

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
17 June 2004 (17.06.2004)

PCT

(10) International Publication Number
WO 2004/050860 A2

(51) International Patent Classification⁷:

C12N

(74) Agents: LICATA, Jane, Massey et al.; Licata & Tyrrell
P.C., 66 East Main Street, Marlton, NJ 08053 (US).

(21) International Application Number:

PCT/US2003/040063

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(22) International Filing Date: 4 December 2003 (04.12.2003)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/431,132 4 December 2002 (04.12.2002) US
60/431,144 4 December 2002 (04.12.2002) US

(84) Designated States (*regional*): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): DI-
ADEXUS, INC. [US/US]; 343 Oyster Point Boulevard,
South San Francisco, CA 94080 (US).

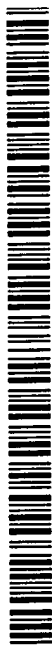
(72) Inventors; and

(75) Inventors/Applicants (*for US only*): MACINA, Roberto,
A. [AR/US]; 4118 Crescendo Avenue, San Jose, CA 95136
(US). TURNER, Leah, R. [US/US]; 939 Rosette Court,
Sunnyvale, CA 94086 (US). SUN, Yongming [CN/US];
551 Shoal Drive, Redwood City, CA 94065 (US).

Published:

— *without international search report and to be republished
upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 2004/050860 A2

(54) Title: COMPOSITIONS, SPLICE VARIANTS AND METHODS RELATING TO COLON SPECIFIC GENES AND PROTEINS

(57) Abstract: The present invention relates to newly identified nucleic acid molecules and polypeptides present in normal and neoplastic colon cells, including fragments, variants and derivatives of the nucleic acids and polypeptides. The present invention also relates to antibodies to the polypeptides of the invention, as well as agonists and antagonists of the polypeptides of the invention. The invention also relates to compositions containing the nucleic acid molecules, polypeptides, antibodies, agonists and antagonists of the invention and methods for the use of these compositions. These uses include identifying, diagnosing, monitoring, staging, imaging and treating colon cancer and non-cancerous disease states in colon, identifying colon tissue, monitoring and identifying and/or designing agonists and antagonists of polypeptides of the invention. The uses also include gene therapy, production of transgenic animals and cells, and production of engineered colon tissue for treatment and research.

10/537402
JC17 Rec'd PCT/PTO 02 JUN 2005

1

COMPOSITIONS, SPLICE VARIANTS AND METHODS RELATING TO COLON SPECIFIC GENES AND PROTEINS

5

INTRODUCTION

This application claims the benefit of priority from U.S. Provisional Patent Application Serial No. 60/431,132 filed December 4, 2002 and 60/431,144 filed December 4, 2002, which are herein incorporated by reference in their entireties.

FIELD OF THE INVENTION

10 The present invention relates to newly identified nucleic acids and polypeptides present in normal and neoplastic colon cells, including fragments, variants and derivatives of the nucleic acids and polypeptides. The present invention also relates to antibodies to the polypeptides of the invention, as well as agonists and antagonists of the polypeptides of the invention. The invention also relates to compositions comprising the nucleic acids,
15 polypeptides, antibodies, post translational modifications (PTMs), variants, derivatives, agonists and antagonists thereof and methods for the use of these compositions. These uses include identifying, diagnosing, monitoring, staging, imaging and treating colon cancer and/or non-cancerous disease states in colon, identifying colon tissue and monitoring and identifying and/or designing agonists and antagonists of polypeptides of
20 the invention. The uses also include gene therapy, therapeutic molecules including but not limited to antibodies or antisense molecules, production of transgenic animals and cells, and production of engineered colon tissue for treatment and research.

BACKGROUND OF THE INVENTION

Colorectal cancer is the second most common cause of cancer death in the United
25 States and the third most prevalent cancer in both men and women. M. L. Davila & A. D. Davila, *Screening for Colon and Rectal Cancer*, in Colon and Rectal Cancer 47 (Peter S. Edelstein ed., 2000). The American Cancer Society estimates that there will be about 105,500 new cases of colon cancer and 42,000 new cases of rectal cancer in 2003 in the United States. Colon cancer and rectal cancer will cause about 57,100 deaths combined.
30 ACS Website: cancer.org on the world wide web. Nearly all cases of colorectal cancer arise from adenomatous polyps, some of which mature into large polyps, undergo abnormal growth and development, and ultimately progress into cancer. Davila at 55-56. This progression would appear to take at least 10 years in most patients, rendering it a

readily treatable form of cancer if diagnosed early, when the cancer is localized. Davila at 56; Walter J. Burdette, Cancer: Etiology, Diagnosis, and Treatment 125 (1998).

Although our understanding of the etiology of colon cancer is undergoing continual refinement, extensive research in this area points to a combination of factors, including age, hereditary and nonhereditary conditions, and environmental/dietary factors. Age is a key risk factor in the development of colorectal cancer, Davila at 48, with men and women over 40 years of age becoming increasingly susceptible to that cancer. Burdette at 126. Incidence rates increase considerably in each subsequent decade of life. Davila at 48. A number of hereditary and nonhereditary conditions have also been linked to a heightened risk of developing colorectal cancer, including familial adenomatous polyposis (FAP), hereditary nonpolyposis colorectal cancer (Lynch syndrome or HNPCC), a personal and/or family history of colorectal cancer or adenomatous polyps, inflammatory bowel disease, diabetes mellitus, and obesity. Davila at 47; Henry T. Lynch & Jane F. Lynch, *Hereditary Nonpolyposis Colorectal Cancer (Lynch Syndromes)*, in Colon and Rectal Cancer 67-68 (Peter S. Edelstein ed., 2000).

Environmental/dietary factors associated with an increased risk of colorectal cancer include a high fat diet, intake of high dietary red meat, and sedentary lifestyle. Davila at 47; Reddy, B. S., *Prev. Med.* 16(4): 460-7 (1987). Conversely, environmental/dietary factors associated with a reduced risk of colorectal cancer include a diet high in fiber, folic acid, calcium, and hormone-replacement therapy in post-menopausal women. Davila at 50-55. The effect of antioxidants in reducing the risk of colon cancer is unclear. Davila at 53.

Because colon cancer is highly treatable when detected at an early, localized stage, screening should be a part of routine care for all adults starting at age 50, especially those with first-degree relatives with colorectal cancer. One major advantage of colorectal cancer screening over its counterparts in other types of cancer is its ability to not only detect precancerous lesions, but to remove them as well. Davila at 56. The key colorectal cancer screening tests in use today are fecal occult blood test, sigmoidoscopy, colonoscopy, double-contrast barium enema, and the carcinoembryonic antigen (CEA) test. Burdette at 125; Davila at 56.

The fecal occult blood test (FOBT) screens for colorectal cancer by detecting the amount of blood in the stool, the premise being that neoplastic tissue, particularly malignant tissue, bleeds more than typical mucosa, with the amount of bleeding increasing

with polyp size and cancer stage. Davila at 56-57. While effective at detecting early stage tumors, FOBT is unable to detect adenomatous polyps (pre-malignant lesions), and, depending on the contents of the fecal sample, is subject to rendering false positives.

5 Davila at 56-59. Sigmoidoscopy and colonoscopy, by contrast, allow direct visualization of the bowel, and enable one to detect, biopsy, and remove adenomatous polyps. Davila at 59-60, 61. Despite the advantages of these procedures, there are accompanying downsides: sigmoidoscopy, by definition, is limited to the sigmoid colon and below, colonoscopy is a relatively expensive procedure, and both share the risk of possible bowel perforation and hemorrhaging. Davila at 59-60. Double-contrast barium enema (DCBE)
10 enables detection of lesions better than FOBT, and almost as well as a colonoscopy, but it may be limited in evaluating the winding rectosigmoid region. Davila at 60. The CEA blood test, which involves screening the blood for carcinoembryonic antigen, shares the downside of FOBT, in that it is of limited utility in detecting colorectal cancer at an early stage. Burdette at 125.

15 Once colon cancer has been diagnosed, treatment decisions are typically made in reference to the stage of cancer progression. A number of techniques are employed to stage the cancer (some of which are also used to screen for colon cancer), including pathologic examination of resected colon, sigmoidoscopy, colonoscopy, and various imaging techniques. AJCC Cancer Staging Handbook 84 (Irvin D. Fleming et al. eds., 5th
20 ed. 1998); Montgomery, R. C. and Ridge, J.A., *Semin. Surg. Oncol.* 15(3): 143-150 (1998). Moreover, chest films, liver functionality tests, and liver scans are employed to determine the extent of metastasis. Fleming at 84. While computerized tomography and magnetic resonance imaging are useful in staging colorectal cancer in its later stages, both have unacceptably low staging accuracy for identifying early stages of the disease, due to
25 the difficulty that both methods have in (1) revealing the depth of bowel wall tumor infiltration and (2) diagnosing malignant adenopathy. Thoeni, R. F., *Radiol. Clin. N. Am.* 35(2): 457-85 (1997). Rather, techniques such as transrectal ultrasound (TRUS) are preferred in this context, although this technique is inaccurate with respect to detecting small lymph nodes that may contain metastases. David Blumberg & Frank G. Opelka,
30 *Neoadjuvant and Adjuvant Therapy for Adenocarcinoma of the Rectum, in Colon and Rectal Cancer* 316 (Peter S. Edelstein ed., 2000).

Several classification systems have been devised to stage the extent of colorectal cancer, including the Dukes' system and the more detailed International Union against

Cancer-American Joint Committee on Cancer TNM staging system, which is considered by many in the field to be a more useful staging system. Burdette at 126-27. The TNM system, which is used for either clinical or pathological staging, is divided into four stages, each of which evaluates the extent of cancer growth with respect to primary tumor (T),
5 regional lymph nodes (N), and distant metastasis (M). Fleming at 84-85. The system focuses on the extent of tumor invasion into the intestinal wall, invasion of adjacent structures, the number of regional lymph nodes that have been affected, and whether distant metastasis has occurred. Fleming at 81.

Stage 0 is characterized by *in situ* carcinoma (Tis), in which the cancer cells are
10 located inside the glandular basement membrane (intraepithelial) or lamina propria (intramucosal). In this stage, the cancer has not spread to the regional lymph nodes (N0), and there is no distant metastasis (M0). In stage I, there is still no spread of the cancer to the regional lymph nodes and no distant metastasis, but the tumor has invaded the submucosa (T1) or has progressed further to invade the muscularis propria (T2). Stage II
15 also involves no spread of the cancer to the regional lymph nodes and no distant metastasis, but the tumor has invaded the subserosa, or the nonperitonealized pericolic or perirectal tissues (T3), or has progressed to invade other organs or structures, and/or has perforated the visceral peritoneum (T4). Stage III is characterized by any of the T substages, no distant metastasis, and either metastasis in 1 to 3 regional lymph nodes (N1)
20 or metastasis in four or more regional lymph nodes (N2). Lastly, stage IV involves any of the T or N substages, as well as distant metastasis. Fleming at 84-85; Burdette at 127.

Currently, pathological staging of colon cancer is preferable over clinical staging as pathological staging provides a more accurate prognosis. Pathological staging typically involves examination of the resected colon section, along with surgical examination of the
25 abdominal cavity. Fleming at 84. Clinical staging would be a preferred method of staging were it at least as accurate as pathological staging, as it does not depend on the invasive procedures of its counterpart.

Turning to the treatment of colorectal cancer, surgical resection results in a cure for roughly 50% of patients. Irradiation is used both preoperatively and postoperatively in
30 treating colorectal cancer. Chemotherapeutic agents, particularly 5-fluorouracil, are also powerful weapons in treating colorectal cancer. Other agents include irinotecan and floxuridine, cisplatin, levamisole, methotrexate, interferon- α , and leucovorin. Burdette at 125, 132-33. Nonetheless, thirty to forty percent of patients will develop a recurrence of

colon cancer following surgical resection, which in many patients is the ultimate cause of death. Wayne De Vos, *Follow-up After Treatment of Colon Cancer*, Colon and Rectal Cancer 225 (Peter S. Edelstein ed., 2000). Accordingly, colon cancer patients must be closely monitored to determine response to therapy and to detect persistent or recurrent disease and metastasis.

The next few paragraphs describe the some of molecular bases of colon cancer. In the case of FAP, the tumor suppressor gene APC (adenomatous polyposis coli), chromosomally located at 5q21, has been either inactivated or deleted by mutation. Alberts et al., Molecular Biology of the Cell 1288 (3d ed. 1994). The APC protein plays a role in a number of functions, including cell adhesion, apoptosis, and repression of the *c-myc* oncogene. N. R. Hall & R. D. Madoff, *Genetics and the Polyp-Cancer Sequence*, Colon and Rectal Cancer 8 (Peter S. Edelstein, ed., 2000). Of those patients with colorectal cancer who have normal APC genes, over 65% have such mutations in the cancer cells but not in other tissues. Alberts et al., *supra* at 1288. In the case of HPNCC, patients manifest abnormalities in the tumor suppressor gene HNPCC, but only about 15% of tumors contain the mutated gene. *Id.* A host of other genes have also been implicated in colorectal cancer, including the *K-ras*, *N-ras*, *H-ras* and *c-myc* oncogenes, and the tumor suppressor genes *DCC* (deleted in colon carcinoma) and *p53*. Hall & Madoff, at 8-9; Alberts et al., at 1288.

Abnormalities in Wg/Wnt signal transduction pathway are also associated with the development of colorectal carcinoma. Taipale, J. and Beachy, P.A. *Nature* 411: 349-354 (2001). Wnt1 is a secreted protein gene originally identified within mouse mammary cancers by its insertion into the mouse mammary tumor virus (MMTV) gene. The protein is homologous to the wingless (Wg) gene product of *Drosophila*, in which it functions as an important factor for the determination of dorsal-ventral segmentation and regulates the formation of fly imaginal discs. Wg/Wnt pathway controls cell proliferation, death and differentiation. Taipal (2001). There are at least 13 members in the Wnt family. These proteins have been found expressed mainly in the central nervous system (CNS) of vertebrates as well as other tissues such as mammary and intestine. The Wnt proteins are the ligands for a family of seven transmembrane domain receptors related to the Frizzled gene product in *Drosophila*. Binding Wnt to Frizzled stimulates the activity of the downstream target, Dishevelled, which in turn inactivates the glycogen synthetase kinase 3β (GSK3 β). Taipal (2001). Usually active GSK3 β will form a complex with the

adenomatous polyposis coli (APC) protein and phosphorylate another complex member, β -catenin. Once phosphorylated, β -catenin is directed to degradation through the ubiquitin pathway. When GSK3 β or APC activity is down regulated, β -catenin is accumulated in the cytoplasm and binds to the T-cell factor or lymphocyte excitation factor (Tcf/Lef) family of transcriptional factors. Binding of β -catenin to Tcf releases the transcriptional repression and induces gene transcription. Among the genes regulated by β -catenin are a transcriptional repressor Engrailed, a transforming growth factor- β (TGF- β) family member Decapentaplegic, and the cytokine Hedgehog in *Drosophila*. β -Catenin is also involved in regulating cell adhesion by binding to α -catenin and E-cadherin. On the other hand, binding of β -catenin to these proteins controls the cytoplasmic β -catenin level and its complexing with TCF. Taipal (2001). Growth factor stimulation and activation of c-src or v-src also regulate β -catenin level by phosphorylation of α -catenin and its related protein, p120^{cas}. When phosphorylated, these proteins decrease their binding to E-cadherin and β -catenin resulting in the accumulation of cytoplasmic β -catenin. Reynolds, A.B. et al. *Mol. Cell Biol.* 14: 8333-8342 (1994). In colon cancer, c-src enzymatic activity has been shown to be increased to the level of v-src. Alternation of components in the Wg/Wnt pathway promotes colorectal carcinoma development. The best known modifications are to the APC gene. Nicola S et al. *Hum. Mol. Genet* 10:721-733 (2001). This germline mutation causes the appearance of hundreds to thousands of adenomatous polyps in the large bowel. It is the gene defect that accounts for the autosomally dominantly inherited FAP and related syndromes. The molecular alternations that occur in this pathway largely involve deletions of alleles of tumor-suppressor genes, such as APC, p53 and Deleted in Colorectal Cancer (DCC), combined with mutational activation of proto-oncogenes, especially c-Ki-ras. Aoki, T. et al. *Human Mutat.* 3: 342-346 (1994). All of these lead to genomic instability in colorectal cancers.

Another source of genomic instability in colorectal cancer is the defect of DNA mismatch repair (MMR) genes. Human homologues of the bacterial *mutHLS* complex (hMSH2, hMLH1, hPMS1, hPMS2 and hMSH6), which is involved in the DNA mismatch repair in bacteria, have been shown to cause the HNPCC (about 70-90% HNPCC) when mutated. Modrich, P. and Lahue, R. *Ann Rev. Biochem.* 65: 101-133 (1996); and Peltomäki, P. *Hum. Mol. Genet* 10: 735-740 (2001). The inactivation of these proteins leads to the accumulation of mutations and causes genetic instability that represents errors

in the accurate replication of the repetitive mono-, di-, tri- and tetra-nucleotide repeats, which are scattered throughout the genome (microsatellite regions). Jass, J.R. et al. *J. Gastroenterol Hepatol* 17: 17-26 (2002). Like in the classic FAP, mutational activation of c-Ki-ras is also required for the promotion of MSI in the alternative HNPCC. Mutations in other proteins such as the tumor suppressor protein phosphatase PTEN (Zhou, X.P. et al. *Hum. Mol. Genet* 11: 445-450 (2002)), BAX (Buttler, L.M. *Aus. N. Z. J. Surg.* 69: 88-94 (1999)), Caspase-5 (Planck, M. *Cancer Genet Cytogenet.* 134: 46-54 (2002)), TGF β -RII (Fallik, D. et al. *Gastroenterol Clin Biol.* 24: 917-22 (2000)) and IGFII-R (Giovannucci E. *J. Nutr.* 131: 3109S-20S (2001)) have also been found in some colorectal tumors possibly as the cause of MMR defect.

Some tyrosine kinases have been shown up-regulated in colorectal tumor tissues or cell lines like HT29. Skoudy, A. et al. *Biochem J.* 317 (Pt 1): 279-84 (1996). Focal adhesion kinase (FAK) and its up-stream kinase c-src and c-yes in colonic epithelial cells may play an important role in the promotion of colorectal cancers through the extracellular matrix (ECM) and integrin-mediated signaling pathways. Jessup, J.M. et al., *The molecular biology of colorectal carcinoma, in: The Molecular Basis of Human Cancer*, 251-268 (Coleman W.B. and Tsongalis G.J. Eds. 2002). The formation of c-src/FAK complexes may coordinately deregulate VEGF expression and apoptosis inhibition. Recent evidences suggest that a specific signal-transduction pathway for cell survival that implicates integrin engagement leads to FAK activation and thus activates PI-3 kinase and akt. In turn, akt phosphorylates BAD and blocks apoptosis in epithelial cells. The activation of c-src in colon cancer may induce VEGF expression through the hypoxia pathway. Other genes that may be implicated in colorectal cancer include Cox enzymes (Ota, S. et al. *Aliment Pharmacol. Ther.* 16 (Suppl 2): 102-106 (2002)), estrogen (al-Azzawi, F. and Wahab, M. *Climacteric* 5: 3-14 (2002)), peroxisome proliferator-activated receptor- γ (PPAR- γ) (Gelman, L. et al. *Cell Mol. Life Sci.* 55: 932-943 (1999)), IGF-I (Giovannucci (2001)), thymine DNA glycosylase (TDG) (Hardeland, U. et al. *Prog. Nucleic Acid Res. Mol. Biol.* 68: 235-253 (2001)) and EGF (Mendelsohn, J. *Endocrine-Related Cancer* 8: 3-9 (2001)).

Gene deletion and mutation are not the only causes for development of colorectal cancers. Epigenetic silencing by DNA methylation also accounts for the loss of function of colorectal cancer suppressor genes. A strong association between MSI and CpG island methylation has been well characterized in sporadic colorectal cancers with high MSI but

not in those of hereditary origin. In one experiment, DNA methylation of MLH1, CDKN2A, MGMT, THBS1, RARB, APC, and p14ARF genes has been shown in 80%, 55%, 23%, 23%, 58%, 35%, and 50% of 40 sporadic colorectal cancers with high MSI respectively. Yamamoto, H. et al. *Genes Chromosomes Cancer* 33: 322-325 (2002); and
5 Kim, K.M. et al. *Oncogene*. 12;21(35): 5441-9 (2002). Carcinogen metabolism enzymes such as GST, NAT, CYP and MTHFR are also associated with an increased or decreased colorectal cancer risk. Pistorius, S. et al. *Kongressbd Dtsch Ges Chir Kongr* 118: 820-824 (2001); and Potter, J.D. *J. Natl. Cancer Inst.* 91: 916-932 (1999).

From the foregoing, it is clear that procedures used for detecting, diagnosing,
10 monitoring, staging, prognosticating, and preventing the recurrence of colorectal cancer are of critical importance to the outcome of the patient. Moreover, current procedures, while helpful in each of these analyses, are limited by their specificity, sensitivity, invasiveness, and/or their cost. As such, highly specific and sensitive procedures that would operate by way of detecting novel markers in cells, tissues, or bodily fluids, with
15 minimal invasiveness and at a reasonable cost, would be highly desirable.

Accordingly, there is a great need for more sensitive and accurate methods for predicting whether a person is likely to develop colorectal cancer, for diagnosing colorectal cancer, for monitoring the progression of the disease, for staging the colorectal cancer, for determining whether the colorectal cancer has metastasized, and for imaging
20 the colorectal cancer. Following accurate diagnosis, there is also a need for less invasive and more effective treatment of colorectal cancer.

Growth and metastasis of solid tumors are also dependent on angiogenesis. Folkman, J., 1986, *Cancer Research*, 46, 467-473; Folkman, J., 1989, *Journal of the National Cancer Institute*, 82, 4-6. It has been shown, for example, that tumors which
25 enlarge to greater than 2 mm must obtain their own blood supply and do so by inducing the growth of new capillary blood vessels. Once these new blood vessels become embedded in the tumor, they provide a means for tumor cells to enter the circulation and metastasize to distant sites such as liver, lung or bone. Weidner, N., et al., 1991, *The New England Journal of Medicine*, 324(1), 1-8.

30 Angiogenesis, defined as the growth or sprouting of new blood vessels from existing vessels, is a complex process that primarily occurs during embryonic development. The process is distinct from vasculogenesis, in that the new endothelial cells lining the vessel arise from proliferation of existing cells, rather than differentiating from

stem cells. The process is invasive and dependent upon proteolysis of the extracellular matrix (ECM), migration of new endothelial cells, and synthesis of new matrix components. Angiogenesis occurs during embryogenic development of the circulatory system; however, in adult humans, angiogenesis only occurs as a response to a
5 pathological condition (except during the reproductive cycle in women).

Under normal physiological conditions in adults, angiogenesis takes place only in very restricted situations such as hair growth and wounding healing. Auerbach, W. and Auerbach, R., 1994, *Pharmacol Ther.* 63(3):265-311; Ribatti et al., 1991, *Haematologica* 76(4):311-20; Risau, 1997, *Nature* 386(6626):671-4. Angiogenesis progresses by a
10 stimulus which results in the formation of a migrating column of endothelial cells. Proteolytic activity is focused at the advancing tip of this "vascular sprout", which breaks down the ECM sufficiently to permit the column of cells to infiltrate and migrate. Behind the advancing front, the endothelial cells differentiate and begin to adhere to each other, thus forming a new basement membrane. The cells then cease proliferation and finally
15 define a lumen for the new arteriole or capillary.

Unregulated angiogenesis has gradually been recognized to be responsible for a wide range of disorders, including, but not limited to, cancer, cardiovascular disease, rheumatoid arthritis, psoriasis and diabetic retinopathy. Folkman, 1995, *Nat Med* 1(1):27-31; Isner, 1999, *Circulation* 99(13):1653-5; Koch, 1998, *Arthritis Rheum* 41(6):951-62;
20 Walsh, 1999, *Rheumatology* (Oxford) 38(2):103-12; Ware and Simons, 1997, *Nat Med* 3(2):158-64.

Of particular interest is the observation that angiogenesis is required by solid tumors for their growth and metastases. Folkman, 1986 *supra*; Folkman 1990, *J Natl. Cancer Inst.*, 82(1) 4-6; Folkman, 1992, *Semin Cancer Biol* 3(2):65-71; Zetter, 1998, *Annu*
25 *Rev Med* 49:407-24. A tumor usually begins as a single aberrant cell which can proliferate only to a size of a few cubic millimeters due to the distance from available capillary beds, and it can stay 'dormant' without further growth and dissemination for a long period of time. Some tumor cells then switch to the angiogenic phenotype to activate endothelial cells, which proliferate and mature into new capillary blood vessels. These newly formed
30 blood vessels not only allow for continued growth of the primary tumor, but also for the dissemination and recolonization of metastatic tumor cells. The precise mechanisms that control the angiogenic switch is not well understood, but it is believed that

neovascularization of tumor mass results from the net balance of a multitude of angiogenesis stimulators and inhibitors Folkman, 1995, *supra*.

One of the most potent angiogenesis inhibitors is endostatin identified by O'Reilly and Folkman. O'Reilly et al., 1997, *Cell* 88(2):277-85; O'Reilly et al., 1994, *Cell* 79(2):3
5 15-28. Its discovery was based on the phenomenon that certain primary tumors can inhibit the growth of distant metastases. O'Reilly and Folkman hypothesized that a primary tumor initiates angiogenesis by generating angiogenic stimulators in excess of inhibitors. However, angiogenic inhibitors, by virtue of their longer half life in the circulation, reach the site of a secondary tumor in excess of the stimulators. The net result is the growth of
10 primary tumor and inhibition of secondary tumor. Endostatin is one of a growing list of such angiogenesis inhibitors produced by primary tumors. It is a proteolytic fragment of a larger protein: endostatin is a 20 kDa fragment of collagen XVIII (amino acid H1132-K1315 in murine collagen XVIII). Endostatin has been shown to specifically inhibit endothelial cell proliferation in vitro and block angiogenesis in vivo. More importantly,
15 administration of endostatin to tumor-bearing mice leads to significant tumor regression, and no toxicity or drug resistance has been observed even after multiple treatment cycles. Boehm et al., 1997, *Nature* 390(6658):404-407. The fact that endostatin targets genetically stable endothelial cells and inhibits a variety of solid tumors makes it a very attractive candidate for anticancer therapy. Fidler and Ellis, 1994, *Cell* 79(2):185-8; Gastl et al.,
20 1997, *Oncology* 54(3):177-84; Hinsbergh et al., 1999, *Ann Oncol* 10 Suppl 4:60-3. In addition, angiogenesis inhibitors have been shown to be more effective when combined with radiation and chemotherapeutic agents. Klement, 2000, *J. Clin Invest*, 105(8) R15-24. Browder, 2000, *Cancer Res.* 6-(7) 1878-86, Arap et al., 1998, *Science* 279(5349):377-80; Mauceri et al., 1998, *Nature* 394(6690):287-91.

25

SUMMARY OF THE INVENTION

The present invention solves many needs in the art by providing nucleic acid molecules, polypeptides and antibodies thereto, variants and derivatives of the nucleic acids and polypeptides, and agonists and antagonists thereto that may be used to identify, diagnose, monitor, stage, image and treat colon cancer and/or non-cancerous disease states
30 in colon; identify and monitor colon tissue; and identify and design agonists and antagonists of polypeptides of the invention. The invention also provides gene therapy,

methods for producing transgenic animals and cells, and methods for producing engineered colon tissue for treatment and research.

One aspect of the present invention relates to nucleic acid molecules that are specific to colon cells, colon tissue and/or the colon organ. These colon specific nucleic acids (CSNAs) may be a naturally occurring cDNA, genomic DNA, RNA, or a fragment
5 of one of these nucleic acids, or may be a non-naturally occurring nucleic acid molecule. If the CSNA is genomic DNA, then the CSNA is a colon specific gene (CSG). If the CSNA is RNA, then it is a colon specific transcript encoded by a CSG. Due to alternative splicing and transcriptional modification one CSG may encode for multiple colon specific
10 RNAs. In a preferred embodiment, the nucleic acid molecule encodes a polypeptide that is specific to colon. More preferred is a nucleic acid molecule that encodes a polypeptide comprising an amino acid sequence of SEQ ID NO: 96-237. In another preferred embodiment, the nucleic acid molecule comprises a nucleic acid sequence of SEQ ID NO: 1-95. For the CSNA sequences listed herein, DEX0448_001.nt.1 corresponds to SEQ ID
15 NO: 1. For sequences with multiple splice variants, the parent sequence DEX0448_001.nt.1, will be followed by DEX0448_001.nt.2, etc. for each splice variant. The sequences of the corresponding peptides are listed as DEX0448_001.aa.1, etc. For the mapping of all of the nucleotides and peptides, see the table in the Example 1 section below.

20 This aspect of the present invention also relates to nucleic acid molecules that selectively hybridize or exhibit substantial sequence similarity to nucleic acid molecules encoding a Colon Specific Protein (CSP), or that selectively hybridize or exhibit substantial sequence similarity to a CSNA. In one embodiment of the present invention the nucleic acid molecule comprises an allelic variant of a nucleic acid molecule encoding
25 a CSP, or an allelic variant of a CSNA. In another embodiment, the nucleic acid molecule comprises a part of a nucleic acid sequence that encodes a CSP or a part of a nucleic acid sequence of a CSNA.

In addition, this aspect of the present invention relates to a nucleic acid molecule further comprising one or more expression control sequences controlling the transcription
30 and/or translation of all or a part of a CSNA or the transcription and/or translation of a nucleic acid molecule that encodes all or a fragment of a CSP.

Another aspect of the present invention relates to vectors and/or host cells comprising a nucleic acid molecule of this invention. In a preferred embodiment, the

nucleic acid molecule of the vector and/or host cell encodes all or a fragment of a CSP. In another preferred embodiment, the nucleic acid molecule of the vector and/or host cell comprises all or a part of a CSNA. Vectors and host cells of the present invention are useful in the recombinant production of polypeptides, particularly CSPs of the present invention.

Another aspect of the present invention relates to polypeptides encoded by a nucleic acid molecule of this invention. The polypeptide may comprise either a fragment or a full-length protein. In a preferred embodiment, the polypeptide is a CSP. However, this aspect of the present invention also relates to mutant proteins (muteins) of CSPs, fusion proteins of which a portion is a CSP, and proteins and polypeptides encoded by allelic variants of a CSNA as provided herein.

A further aspect of the present invention is a novel splice variant which encodes an amino acid sequence that provides a novel region to be targeted for the generation of reagents that can be used in the detection and/or treatment of cancer. The novel amino acid sequence may lead to a unique protein structure, protein subcellular localization, biochemical processing or function. This information can be used to directly or indirectly facilitate the generation of additional or novel therapeutics or diagnostics. The nucleotide sequence in this novel splice variant can be used as a nucleic acid probe for the diagnosis and/or treatment of cancer.

Another aspect of the present invention relates to antibodies and other binders that specifically bind to a polypeptide of the instant invention. Accordingly antibodies or binders of the present invention specifically bind to CSPs, muteins, fusion proteins, and/or homologous proteins or polypeptides encoded by allelic variants of a CSNA as provided herein.

Another aspect of the present invention relates to agonists and antagonists of the nucleic acid molecules and polypeptides of this invention. The agonists and antagonists of the instant invention may be used to treat colon cancer and non-cancerous disease states in colon and to produce engineered colon tissue.

Another aspect of the present invention relates to methods for using the nucleic acid molecules to detect or amplify nucleic acid molecules that have similar or identical nucleic acid sequences compared to the nucleic acid molecules described herein. Such methods are useful in identifying, diagnosing, monitoring, staging, imaging and treating colon cancer and/or non-cancerous disease states in colon. Such methods are also useful

in identifying and/or monitoring colon tissue. In addition, measurement of levels of one or more of the nucleic acid molecules of this invention may be useful as a diagnostic as part of a panel in combination with known other markers, particularly those described in the colon cancer background section above.

- 5 Another aspect of the present invention relates to use of the nucleic acid molecules of this invention in gene therapy, for producing transgenic animals and cells, and for producing engineered colon tissue for treatment and research.

- Another aspect of the present invention relates to methods for detecting polypeptides of this invention, preferably using antibodies thereto. Such methods are
10 useful to identify, diagnose, monitor, stage, image and treat colon cancer and non-cancerous disease states in colon. In addition, measurement of levels of one or more of the polypeptides of this invention may be useful to identify, diagnose, monitor, stage, and/or image colon cancer in combination with known other markers, particularly those described in the colon cancer background section above. The polypeptides of the present
15 invention can also be used to identify and/or monitor colon tissue, and to produce engineered colon tissue.

- Yet another aspect of the present invention relates to a computer readable means of storing the nucleic acid and amino acid sequences of the invention. The records of the computer readable means can be accessed for reading and displaying of sequences for
20 comparison, alignment and ordering of the sequences of the invention to other sequences. In addition, the computer records regarding the nucleic acid and/or amino acid sequences and/or measurements of their levels may be used alone or in combination with other markers to diagnose colon related diseases.

DETAILED DESCRIPTION OF THE INVENTION

25 Definitions and General Techniques

- Unless otherwise defined herein, scientific and technical terms used in connection with the present invention shall have the meanings that are commonly understood by those of ordinary skill in the art. Further, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular. Generally,
30 nomenclatures used in connection with, and techniques of, cell and tissue culture, molecular biology, immunology, microbiology, genetics and protein and nucleic acid chemistry and hybridization described herein are those well known and commonly used in

the art. The methods and techniques of the present invention are generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification unless otherwise indicated. *See, e.g.,* Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press (1989) and Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 3d ed., Cold Spring Harbor Press (2001); Ausubel *et al.*, Current Protocols in Molecular Biology, Greene Publishing Associates (1992, and Supplements to 2000); Ausubel *et al.*, Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology – 4th Ed., Wiley & Sons (1999); Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press (1990); and Harlow and Lane, Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press (1999).

Enzymatic reactions and purification techniques are performed according to manufacturer's specifications, as commonly accomplished in the art or as described herein. The nomenclatures used in connection with, and the laboratory procedures and techniques of, analytical chemistry, synthetic organic chemistry, and medicinal and pharmaceutical chemistry described herein are those well known and commonly used in the art. Standard techniques are used for chemical syntheses, chemical analyses, pharmaceutical preparation, formulation, and delivery, and treatment of patients.

The following terms, unless otherwise indicated, shall be understood to have the following meanings:

A "nucleic acid molecule" of this invention refers to a polymeric form of nucleotides and includes both sense and antisense strands of RNA, cDNA, genomic DNA, and synthetic forms and mixed polymers of the above. A nucleotide refers to a ribonucleotide, deoxynucleotide or a modified form of either type of nucleotide. A "nucleic acid molecule" as used herein is synonymous with "nucleic acid" and "polynucleotide." The term "nucleic acid molecule" usually refers to a molecule of at least 10 bases in length, unless otherwise specified. The term includes single- and double-stranded forms of DNA. In addition, a polynucleotide may include either or both naturally occurring and modified nucleotides linked together by naturally occurring and/or non-naturally occurring nucleotide linkages.

Nucleotides are represented by single letter symbols in nucleic acid molecule sequences. The following table lists symbols identifying nucleotides or groups of

nucleotides which may occupy the symbol position on a nucleic acid molecule. See Nomenclature Committee of the International Union of Biochemistry (NC-IUB), Nomenclature for incompletely specified bases in nucleic acid sequences, Recommendations 1984., *Eur J Biochem.* 150(1):1-5 (1985).

Symbol	Meaning	Group/Origin of Designation	Complementary Symbol
a	a	Adenine	t/u
g	g	Guanine	c
c	c	Cytosine	g
t	t	Thymine	a
u	u	Uracil	a
r	g or a	puRine	y
y	t/u or c	pYrimidine	r
m	a or c	aMino	k
k	g or t/u	Keto	m
s	g or c	Strong interactions 3H-bonds	w
w	a or t/u	Weak interactions 2H-bonds	s
b	g or c or t/u	not a	v
d	a or g or t/u	not c	h
h	a or c or t/u	not g	d
v	a or g or c	not t, not u	b
n	a or g or c or t/u, unknown, or other	aNy	n

5

The nucleic acid molecules may be modified chemically or biochemically or may contain non-natural or derivatized nucleotide bases, as will be readily appreciated by those of skill in the art. Such modifications include, for example, labels, methylation, substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as uncharged linkages (*e.g.*, methyl phosphonates, phosphotriesters, phosphoramidates, carbamates, etc.), charged linkages (*e.g.*, phosphorothioates, phosphorodithioates, etc.), pendent moieties (*e.g.*, polypeptides), intercalators (*e.g.*, acridine, psoralen, etc.), chelators, alkylators, and modified linkages (*e.g.*, alpha anomeric nucleic acids, etc.) The term "nucleic acid molecule" also includes any topological conformation, including single-stranded, double-stranded, partially duplexed, triplexed, hairpinned, circular and padlocked conformations. Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions. Such molecules are known in the art and include, for example, those in which peptide linkages substitute for phosphate linkages in the backbone of the molecule.

20

A "gene" is defined as a nucleic acid molecule that comprises a nucleic acid sequence that encodes a polypeptide and the expression control sequences that surround the nucleic acid sequence that encodes the polypeptide. For instance, a gene may comprise a promoter, one or more enhancers, a nucleic acid sequence that encodes a polypeptide, downstream regulatory sequences and, possibly, other nucleic acid sequences involved in regulation of the expression of an RNA. As is well known in the art, eukaryotic genes usually contain both exons and introns. The term "exon" refers to a nucleic acid sequence found in genomic DNA that is bioinformatically predicted and/or experimentally confirmed to contribute contiguous sequence to a mature mRNA transcript. The term "intron" refers to a nucleic acid sequence found in genomic DNA that is predicted and/or confirmed to not contribute to a mature mRNA transcript, but rather to be "spliced out" during processing of the transcript.

A nucleic acid molecule or polypeptide is "derived" from a particular species if the nucleic acid molecule or polypeptide has been isolated from the particular species, or if the nucleic acid molecule or polypeptide is homologous to a nucleic acid molecule or polypeptide isolated from a particular species.

An "isolated" or "substantially pure" nucleic acid or polynucleotide (*e.g.*, an RNA, DNA or a mixed polymer) is one which is substantially separated from other cellular components that naturally accompany the native polynucleotide in its natural host cell, *e.g.*, ribosomes, polymerases, or genomic sequences with which it is naturally associated. The term embraces a nucleic acid or polynucleotide that (1) has been removed from its naturally occurring environment, (2) is not associated with all or a portion of a polynucleotide in which the "isolated polynucleotide" is found in nature, (3) is operatively linked to a polynucleotide which it is not linked to in nature, (4) does not occur in nature as part of a larger sequence or (5) includes nucleotides or internucleoside bonds that are not found in nature. The term "isolated" or "substantially pure" also can be used in reference to recombinant or cloned DNA isolates, chemically synthesized polynucleotide analogs, or polynucleotide analogs that are biologically synthesized by heterologous systems. The term "isolated nucleic acid molecule" includes nucleic acid molecules that are integrated into a host cell chromosome at a heterologous site, recombinant fusions of a native fragment to a heterologous sequence, recombinant vectors present as episomes or as integrated into a host cell chromosome.

A "part" of a nucleic acid molecule refers to a nucleic acid molecule that comprises a partial contiguous sequence of at least 10 bases of the reference nucleic acid molecule. Preferably, a part comprises at least 15 to 20 bases of a reference nucleic acid molecule. In theory, a nucleic acid sequence of 17 nucleotides is of sufficient length to
5 occur at random less frequently than once in the three gigabase human genome, and thus provides a nucleic acid probe that can uniquely identify the reference sequence in a nucleic acid mixture of genomic complexity. A preferred part is one that comprises a nucleic acid sequence that can encode at least 6 contiguous amino acid sequences (fragments of at least 18 nucleotides) because they are useful in directing the expression or
10 synthesis of peptides that are useful in mapping the epitopes of the polypeptide encoded by the reference nucleic acid. *See, e.g., Geysen et al., Proc. Natl. Acad. Sci. USA* 81:3998-4002 (1984); and U.S. Patent Nos. 4,708,871 and 5,595,915, the disclosures of which are incorporated herein by reference in their entireties. A part may also comprise at least 25, 30, 35 or 40 nucleotides of a reference nucleic acid molecule, or at least 50, 60,
15 70, 80, 90, 100, 150, 200, 250, 300, 350, 400 or 500 nucleotides of a reference nucleic acid molecule. A part of a nucleic acid molecule may comprise no other nucleic acid sequences. Alternatively, a part of a nucleic acid may comprise other nucleic acid sequences from other nucleic acid molecules.

The term "oligonucleotide" refers to a nucleic acid molecule generally comprising
20 a length of 200 bases or fewer. The term often refers to single-stranded deoxyribonucleotides, but it can refer as well to single-or double-stranded ribonucleotides, RNA:DNA hybrids and double-stranded DNAs, among others. Preferably, oligonucleotides are 10 to 60 bases in length and most preferably 12, 13, 14, 15, 16, 17, 18, 19 or 20 bases in length. Other preferred oligonucleotides are 25, 30, 35, 40, 45, 50,
25 55 or 60 bases in length. Oligonucleotides may be single-stranded, *e.g.* for use as probes or primers, or may be double-stranded, *e.g.* for use in the construction of a mutant gene. Oligonucleotides of the invention can be either sense or antisense oligonucleotides. An oligonucleotide can be derivatized or modified as discussed above for nucleic acid molecules.

30 Oligonucleotides, such as single-stranded DNA probe oligonucleotides, often are synthesized by chemical methods, such as those implemented on automated oligonucleotide synthesizers. However, oligonucleotides can be made by a variety of other methods, including in vitro recombinant DNA-mediated techniques and by

expression of DNAs in cells and organisms. Initially, chemically synthesized DNAs typically are obtained without a 5' phosphate. The 5' ends of such oligonucleotides are not substrates for phosphodiester bond formation by ligation reactions that employ DNA ligases typically used to form recombinant DNA molecules. Where ligation of such
5 oligonucleotides is desired, a phosphate can be added by standard techniques, such as those that employ a kinase and ATP. The 3' end of a chemically synthesized oligonucleotide generally has a free hydroxyl group and, in the presence of a ligase, such as T4 DNA ligase, readily will form a phosphodiester bond with a 5' phosphate of another polynucleotide, such as another oligonucleotide. As is well known, this reaction can be
10 prevented selectively, where desired, by removing the 5' phosphates of the other polynucleotide(s) prior to ligation.

The term "naturally occurring nucleotide" referred to herein includes naturally occurring deoxyribonucleotides and ribonucleotides. The term "modified nucleotides" referred to herein includes nucleotides with modified or substituted sugar groups and the
15 like. The term "nucleotide linkages" referred to herein includes nucleotide linkages such as phosphorothioate, phosphorodithioate, phosphoroselenoate, phosphorodiselenoate, phosphoroanilothioate, phosphoraniladate, phosphoroamidate, and the like. *See e.g.*, LaPlanche *et al. Nucl. Acids Res.* 14:9081-9093 (1986); Stein *et al. Nucl. Acids Res.* 16:3209-3221 (1988); Zon *et al. Anti-Cancer Drug Design* 6:539-568 (1991); Zon *et al.*,
20 in Eckstein (ed.) Oligonucleotides and Analogues: A Practical Approach, pp. 87-108, Oxford University Press (1991); Uhlmann and Peyman *Chemical Reviews* 90:543 (1990), and U.S. Patent No. 5,151,510, the disclosure of which is hereby incorporated by reference in its entirety.

Unless specified otherwise, the left hand end of a polynucleotide sequence in sense
25 orientation is the 5' end and the right hand end of the sequence is the 3' end. In addition, the left hand direction of a polynucleotide sequence in sense orientation is referred to as the 5' direction, while the right hand direction of the polynucleotide sequence is referred to as the 3' direction. Further, unless otherwise indicated, each nucleotide sequence is set forth herein as a sequence of deoxyribonucleotides. It is intended, however, that the given
30 sequence be interpreted as would be appropriate to the polynucleotide composition: for example, if the isolated nucleic acid is composed of RNA, the given sequence intends ribonucleotides, with uridine substituted for thymidine.

The term “allelic variant” refers to one of two or more alternative naturally occurring forms of a gene, wherein each gene possesses a unique nucleotide sequence. In a preferred embodiment, different alleles of a given gene have similar or identical biological properties.

5 The term “percent sequence identity” in the context of nucleic acid sequences refers to the residues in two sequences which are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over a stretch of at least about nine nucleotides, usually at least about 20 nucleotides, more usually at least about 24 nucleotides, typically at least about 28 nucleotides, more typically at least about
10 32 nucleotides, and preferably at least about 36 or more nucleotides. There are a number of different algorithms known in the art which can be used to measure nucleotide sequence identity. For instance, polynucleotide sequences can be compared using FASTA, Gap or Bestfit, which are programs in Wisconsin Package Version 10.0, Genetics Computer
15 Group (GCG), Madison, Wisconsin. FASTA, which includes, *e.g.*, the programs FASTA2 and FASTA3, provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences (Pearson, *Methods Enzymol.* 183: 63-98 (1990); Pearson, *Methods Mol. Biol.* 132: 185-219 (2000); Pearson, *Methods Enzymol.* 266: 227-258 (1996); Pearson, *J. Mol. Biol.* 276: 71-84 (1998)). Unless otherwise specified, default parameters for a particular program or algorithm are used. For instance,
20 percent sequence identity between nucleic acid sequences can be determined using FASTA with its default parameters (a word size of 6 and the NOPAM factor for the scoring matrix) or using Gap with its default parameters as provided in GCG Version 6.1.

A reference to a nucleic acid sequence encompasses its complement unless otherwise specified. Thus, a reference to a nucleic acid molecule having a particular
25 sequence should be understood to encompass its complementary strand, with its complementary sequence. The complementary strand is also useful, *e.g.*, for antisense therapy, double-stranded RNA (dsRNA) inhibition (RNAi), combination of triplex and antisense, hybridization probes and PCR primers.

In the molecular biology art, researchers use the terms “percent sequence identity”,
30 “percent sequence similarity” and “percent sequence homology” interchangeably. In this application, these terms shall have the same meaning with respect to nucleic acid sequences only.

The term “substantial similarity” or “substantial sequence similarity,” when referring to a nucleic acid or fragment thereof, indicates that, when optimally aligned with appropriate nucleotide insertions or deletions with another nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 50%, more preferably 60% of the nucleotide bases, usually at least about 70%, more usually at least about 80%, preferably at least about 90%, and more preferably at least about 95-98% of the nucleotide bases, as measured by any well known algorithm of sequence identity, such as FASTA, BLAST or Gap, as discussed above.

Alternatively, substantial similarity exists between a first and second nucleic acid sequence when the first nucleic acid sequence or fragment thereof hybridizes to an antisense strand of the second nucleic acid, under selective hybridization conditions. Typically, selective hybridization will occur between the first nucleic acid sequence and an antisense strand of the second nucleic acid sequence when there is at least about 55% sequence identity between the first and second nucleic acid sequences—preferably at least about 65%, more preferably at least about 75%, and most preferably at least about 90% — over a stretch of at least about 14 nucleotides, more preferably at least 17 nucleotides, even more preferably at least 20, 25, 30, 35, 40, 50, 60, 70, 80, 90 or 100 nucleotides.

Nucleic acid hybridization will be affected by such conditions as salt concentration, temperature, solvents, the base composition of the hybridizing species, length of the complementary regions, and the number of nucleotide base mismatches between the hybridizing nucleic acids, as will be readily appreciated by those skilled in the art. “Stringent hybridization conditions” and “stringent wash conditions” in the context of nucleic acid hybridization experiments depend upon a number of different physical parameters. The most important parameters include temperature of hybridization, base composition of the nucleic acids, salt concentration and length of the nucleic acid. One having ordinary skill in the art knows how to vary these parameters to achieve a particular stringency of hybridization. In general, “stringent hybridization” is performed at about 25°C below the thermal melting point (T_m) for the specific DNA hybrid under a particular set of conditions. “Stringent washing” is performed at temperatures about 5°C lower than the T_m for the specific DNA hybrid under a particular set of conditions. The T_m is the temperature at which 50% of the target sequence hybridizes to a perfectly matched probe. See Sambrook (1989), *supra*, p. 9.51.

The T_m for a particular DNA-DNA hybrid can be estimated by the formula:

$T_m = 81.5^{\circ}\text{C} + 16.6 (\log_{10}[\text{Na}^+]) + 0.41 (\text{fraction G} + \text{C}) - 0.63 (\% \text{ formamide}) - (600/l)$ where l is the length of the hybrid in base pairs.

The T_m for a particular RNA-RNA hybrid can be estimated by the formula:

$T_m = 79.8^{\circ}\text{C} + 18.5 (\log_{10}[\text{Na}^+]) + 0.58 (\text{fraction G} + \text{C}) +$

5 $11.8 (\text{fraction G} + \text{C})^2 - 0.35 (\% \text{ formamide}) - (820/l).$

The T_m for a particular RNA-DNA hybrid can be estimated by the formula:

$T_m = 79.8^{\circ}\text{C} + 18.5 (\log_{10}[\text{Na}^+]) + 0.58 (\text{fraction G} + \text{C}) +$

$11.8 (\text{fraction G} + \text{C})^2 - 0.50 (\% \text{ formamide}) - (820/l).$

In general, the T_m decreases by 1-1.5°C for each 1% of mismatch between two
 10 nucleic acid sequences. Thus, one having ordinary skill in the art can alter hybridization and/or washing conditions to obtain sequences that have higher or lower degrees of sequence identity to the target nucleic acid. For instance, to obtain hybridizing nucleic acids that contain up to 10% mismatch from the target nucleic acid sequence, 10-15°C would be subtracted from the calculated T_m of a perfectly matched hybrid, and then the
 15 hybridization and washing temperatures adjusted accordingly. Probe sequences may also hybridize specifically to duplex DNA under certain conditions to form triplex or other higher order DNA complexes. The preparation of such probes and suitable hybridization conditions are well known in the art.

An example of stringent hybridization conditions for hybridization of
 20 complementary nucleic acid sequences having more than 100 complementary residues on a filter in a Southern or Northern blot or for screening a library is 50% formamide/6X SSC at 42°C for at least ten hours and preferably overnight (approximately 16 hours). Another example of stringent hybridization conditions is 6X SSC at 68°C without formamide for at least ten hours and preferably overnight. An example of moderate stringency
 25 hybridization conditions is 6X SSC at 55°C without formamide for at least ten hours and preferably overnight. An example of low stringency hybridization conditions for hybridization of complementary nucleic acid sequences having more than 100 complementary residues on a filter in a Southern or northern blot or for screening a library is 6X SSC at 42°C for at least ten hours. Hybridization conditions to identify nucleic acid
 30 sequences that are similar but not identical can be identified by experimentally changing the hybridization temperature from 68°C to 42°C while keeping the salt concentration constant (6X SSC), or keeping the hybridization temperature and salt concentration constant (e.g. 42°C and 6X SSC) and varying the formamide concentration from 50% to

0%. Hybridization buffers may also include blocking agents to lower background. These agents are well known in the art. *See* Sambrook *et al.* (1989), *supra*, pages 8.46 and 9.46-9.58. *See also* Ausubel (1992), *supra*, Ausubel (1999), *supra*, and Sambrook (2001), *supra*.

5 Wash conditions also can be altered to change stringency conditions. An example of stringent wash conditions is a 0.2x SSC wash at 65°C for 15 minutes (*see* Sambrook (1989), *supra*, for SSC buffer). Often the high stringency wash is preceded by a low stringency wash to remove excess probe. An exemplary medium stringency wash for duplex DNA of more than 100 base pairs is 1x SSC at 45°C for 15 minutes. An
10 exemplary low stringency wash for such a duplex is 4x SSC at 40°C for 15 minutes. In general, signal-to-noise ratio of 2x or higher than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization.

As defined herein, nucleic acids that do not hybridize to each other under stringent conditions are still substantially similar to one another if they encode polypeptides that are
15 substantially identical to each other. This occurs, for example, when a nucleic acid is created synthetically or recombinantly using a high codon degeneracy as permitted by the redundancy of the genetic code.

Hybridization conditions for nucleic acid molecules that are shorter than 100 nucleotides in length (e.g., for oligonucleotide probes) may be calculated by the formula:
20
$$T_m = 81.5^{\circ}\text{C} + 16.6(\log_{10}[\text{Na}^+]) + 0.41(\text{fraction G+C}) - (600/\text{N})$$
, wherein N is change length and the $[\text{Na}^+]$ is 1 M or less. *See* Sambrook (1989), *supra*, p. 11.46. For hybridization of probes shorter than 100 nucleotides, hybridization is usually performed under stringent conditions (5-10°C below the T_m) using high concentrations (0.1-1.0 pmol/ml) of probe. *Id.* at p. 11.45. Determination of hybridization using mismatched
25 probes, pools of degenerate probes or “guessmers,” as well as hybridization solutions and methods for empirically determining hybridization conditions are well known in the art. *See, e.g.*, Ausubel (1999), *supra*; Sambrook (1989), *supra*, pp. 11.45-11.57.

The term “digestion” or “digestion of DNA” refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The
30 various restriction enzymes referred to herein are commercially available and their reaction conditions, cofactors and other requirements for use are known and routine to the skilled artisan. For analytical purposes, typically, 1 µg of plasmid or DNA fragment is digested with about 2 units of enzyme in about 20 µl of reaction buffer. For the purpose of

isolating DNA fragments for plasmid construction, typically 5 to 50 µg of DNA are digested with 20 to 250 units of enzyme in proportionately larger volumes. Appropriate buffers and substrate amounts for particular restriction enzymes are described in standard laboratory manuals, such as those referenced below, and are specified by commercial suppliers. Incubation times of about 1 hour at 37°C are ordinarily used, but conditions may vary in accordance with standard procedures, the supplier's instructions and the particulars of the reaction. After digestion, reactions may be analyzed, and fragments may be purified by electrophoresis through an agarose or polyacrylamide gel, using well known methods that are routine for those skilled in the art.

10 The term "ligation" refers to the process of forming phosphodiester bonds between two or more polynucleotides, which most often are double-stranded DNAs. Techniques for ligation are well known to the art and protocols for ligation are described in standard laboratory manuals and references, such as, *e.g.*, Sambrook (1989), *supra*.

Genome-derived "single exon probes," are probes that comprise at least part of an exon ("reference exon") and can hybridize detectably under high stringency conditions to transcript-derived nucleic acids that include the reference exon but do not hybridize detectably under high stringency conditions to nucleic acids that lack the reference exon. Single exon probes typically further comprise, contiguous to a first end of the exon portion, a first intronic and/or intergenic sequence that is identically contiguous to the exon in the genome, and may contain a second intronic and/or intergenic sequence that is identically contiguous to the exon in the genome. The minimum length of genome-derived single exon probes is defined by the requirement that the exonic portion be of sufficient length to hybridize under high stringency conditions to transcript-derived nucleic acids, as discussed above. The maximum length of genome-derived single exon probes is defined by the requirement that the probes contain portions of no more than one exon. The single exon probes may contain priming sequences not found in contiguity with the rest of the probe sequence in the genome, which priming sequences are useful for PCR and other amplification-based technologies. In another aspect, the invention is directed to single exon probes based on the CSNAs disclosed herein.

30 In one embodiment, the term "microarray" refers to a "nucleic acid microarray" having a substrate-bound plurality of nucleic acids, hybridization to each of the plurality of bound nucleic acids being separately detectable. The substrate can be solid or porous, planar or non-planar, unitary or distributed. Nucleic acid microarrays include all the

devices so called in Schena (ed.), DNA Microarrays: A Practical Approach (Practical Approach Series), Oxford University Press (1999); *Nature Genet.* 21(1)(suppl.):1 - 60 (1999); Schena (ed.), Microarray Biochip: Tools and Technology, Eaton Publishing Company/BioTechniques Books Division (2000). Additionally, these nucleic acid

5 microarrays include a substrate-bound plurality of nucleic acids in which the plurality of nucleic acids are disposed on a plurality of beads, rather than on a unitary planar substrate, as is described, *inter alia*, in Brenner *et al.*, *Proc. Natl. Acad. Sci. USA* 97(4):1665-1670 (2000). Examples of nucleic acid microarrays may be found in U.S. Patent Nos.

6,391,623, 6,383,754, 6,383,749, 6,380,377, 6,379,897, 6,376,191, 6,372,431, 6,351,712

10 6,344,316, 6,316,193, 6,312,906, 6,309,828, 6,309,824, 6,306,643, 6,300,063, 6,287,850, 6,284,497, 6,284,465, 6,280,954, 6,262,216, 6,251,601, 6,245,518, 6,263,287, 6,251,601, 6,238,866, 6,228,575, 6,214,587, 6,203,989, 6,171,797, 6,103,474, 6,083,726, 6,054,274, 6,040,138, 6,083,726, 6,004,755, 6,001,309, 5,958,342, 5,952,180, 5,936,731, 5,843,655, 5,814,454, 5,837,196, 5,436,327, 5,412,087, and 5,405,783, the disclosures of which are

15 incorporated herein by reference in their entirety.

In an alternative embodiment, a "microarray" may also refer to a "peptide microarray" or "protein microarray" having a substrate-bound collection or plurality of polypeptides, the binding to each of the plurality of bound polypeptides being separately detectable. Alternatively, the peptide microarray may have a plurality of binders,

20 including but not limited to monoclonal antibodies, polyclonal antibodies, phage display binders, yeast 2 hybrid binders, and aptamers, which can specifically detect the binding of the polypeptides of this invention. The array may be based on autoantibody detection to the polypeptides of this invention, see Robinson *et al.*, *Nature Medicine* 8(3):295-301 (2002). Examples of peptide arrays may be found in WO 02/31463, WO 02/25288, WO

25 01/94946, WO 01/88162, WO 01/68671, WO 01/57259, WO 00/61806, WO 00/54046, WO 00/47774, WO 99/40434, WO 99/39210, and WO 97/42507 and U.S. Patent Nos. 6,268,210, 5,766,960, and 5,143,854, the disclosures of which are incorporated herein by reference in their entirety.

In addition, determination of the levels of the CSNA or CSP may be made in a

30 multiplex manner using techniques described in WO 02/29109, WO 02/24959, WO 01/83502, WO01/73113, WO 01/59432, WO 01/57269, and WO 99/67641, the disclosures of which are incorporated herein by reference in their entirety.

The term "mutant", "mutated", or "mutation" when applied to nucleic acid sequences means that nucleotides in a nucleic acid sequence may be inserted, deleted or changed compared to a reference nucleic acid sequence. A single alteration may be made at a locus (a point mutation) or multiple nucleotides may be inserted, deleted or changed at a single locus. In addition, one or more alterations may be made at any number of loci within a nucleic acid sequence. In a preferred embodiment of the present invention, the nucleic acid sequence is the wild type nucleic acid sequence encoding a CSP or is a CSNA. The nucleic acid sequence may be mutated by any method known in the art including those mutagenesis techniques described *infra*.

10 The term "error-prone PCR" refers to a process for performing PCR under conditions where the copying fidelity of the DNA polymerase is low, such that a high rate of point mutations is obtained along the entire length of the PCR product. *See, e.g., Leung et al., Technique 1: 11-15 (1989) and Caldwell et al., PCR Methods Applic. 2: 28-33 (1992).*

15 The term "oligonucleotide-directed mutagenesis" refers to a process which enables the generation of site-specific mutations in any cloned DNA segment of interest. *See, e.g., Reidhaar-Olson et al., Science 241: 53-57 (1988).*

The term "assembly PCR" refers to a process which involves the assembly of a PCR product from a mixture of small DNA fragments. A large number of different PCR reactions occur in parallel in the same vial, with the products of one reaction priming the products of another reaction.

20 The term "sexual PCR mutagenesis" or "DNA shuffling" refers to a method of error-prone PCR coupled with forced homologous recombination between DNA molecules of different but highly related DNA sequence *in vitro*, caused by random fragmentation of the DNA molecule based on sequence similarity, followed by fixation of the crossover by primer extension in an error-prone PCR reaction. *See, e.g., Stemmer, Proc. Natl. Acad. Sci. U.S.A. 91: 10747-10751 (1994).* DNA shuffling can be carried out between several related genes ("Family shuffling").

25 The term "*in vivo* mutagenesis" refers to a process of generating random mutations in any cloned DNA of interest which involves the propagation of the DNA in a strain of bacteria such as *E. coli* that carries mutations in one or more of the DNA repair pathways. These "mutator" strains have a higher random mutation rate than that of a wild-type

parent. Propagating the DNA in a mutator strain will eventually generate random mutations within the DNA.

The term "cassette mutagenesis" refers to any process for replacing a small region of a double-stranded DNA molecule with a synthetic oligonucleotide "cassette" that
5 differs from the native sequence. The oligonucleotide often contains completely and/or partially randomized native sequence.

The term "recursive ensemble mutagenesis" refers to an algorithm for protein engineering (protein mutagenesis) developed to produce diverse populations of phenotypically related mutants whose members differ in amino acid sequence. This
10 method uses a feedback mechanism to control successive rounds of combinatorial cassette mutagenesis. *See, e.g., Arkin et al., Proc. Natl. Acad. Sci. U.S.A.* 89: 7811-7815 (1992).

The term "exponential ensemble mutagenesis" refers to a process for generating combinatorial libraries with a high percentage of unique and functional mutants, wherein small groups of residues are randomized in parallel to identify, at each altered position,
15 amino acids which lead to functional proteins. *See, e.g., Delegrave et al., Biotechnology Research* 11: 1548-1552 (1993); Arnold, *Current Opinion in Biotechnology* 4: 450-455 (1993).

"Operatively linked" expression control sequences refers to a linkage in which the expression control sequence is either contiguous with the gene of interest to control the
20 gene of interest, or acts in *trans* or at a distance to control the gene of interest.

The term "expression control sequence" as used herein refers to polynucleotide sequences which are necessary to affect the expression of coding sequences to which they are operatively linked. Expression control sequences are sequences which control the transcription, post-transcriptional events and translation of nucleic acid sequences.
25 Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (*e.g.*, ribosome binding sites); sequences that enhance protein stability; and when desired, sequences that enhance protein secretion. The nature
30 of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include promoter, ribosomal binding site, and transcription termination sequence. The term "control sequences" is intended to include, at a minimum, all components whose presence is essential for expression, and can also include additional

components whose presence is advantageous, for example, leader sequences and fusion partner sequences.

The term "vector," as used herein, is intended to refer to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double-stranded DNA loop into which additional DNA segments may be ligated. Other vectors include cosmids, bacterial artificial chromosomes (BAC) and yeast artificial chromosomes (YAC). Another type of vector is a viral vector, wherein additional DNA segments may be ligated into the viral genome. Viral vectors that infect bacterial cells are referred to as bacteriophages. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication). Other vectors can be integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "recombinant expression vectors" (or simply, "expression vectors"). In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" may be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include other forms of expression vectors that serve equivalent functions.

The term "recombinant host cell" (or simply "host cell"), as used herein, is intended to refer to a cell into which a recombinant expression vector has been introduced. It should be understood that such terms are intended to refer not only to the particular subject cell but also to the progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term "host cell" as used herein.

As used herein, the phrase "open reading frame" and the equivalent acronym "ORF" refers to that portion of a transcript-derived nucleic acid that can be translated in its entirety into a sequence of contiguous amino acids. As so defined, an ORF has length, measured in nucleotides, exactly divisible by 3. As so defined, an ORF need not encode the entirety of a natural protein.

As used herein, the phrase "ORF-encoded peptide" refers to the predicted or actual translation of an ORF.

As used herein, the phrase "degenerate variant" of a reference nucleic acid sequence is meant to be inclusive of all nucleic acid sequences that can be directly translated, using the standard genetic code, to provide an amino acid sequence identical to that translated from the reference nucleic acid sequence.

The term "polypeptide" encompasses both naturally occurring and non-naturally occurring proteins and polypeptides, as well as polypeptide fragments and polypeptide mutants, derivatives and analogs thereof. A polypeptide may be monomeric or polymeric. Further, a polypeptide may comprise a number of different modules within a single polypeptide each of which has one or more distinct activities. A preferred polypeptide in accordance with the invention comprises a CSP encoded by a nucleic acid molecule of the instant invention, or a fragment, mutant, analog or derivative thereof.

The term "isolated protein" or "isolated polypeptide" is a protein or polypeptide that by virtue of its origin or source of derivation (1) is not associated with naturally associated components that accompany it in its native state, (2) is free of other proteins from the same species (3) is expressed by a cell from a different species, or (4) does not occur in nature. Thus, a polypeptide that is chemically synthesized or synthesized in a cellular system different from the cell from which it naturally originates will be "isolated" from its naturally associated components. A polypeptide or protein may also be rendered substantially free of naturally associated components by isolation, using protein purification techniques well known in the art.

A protein or polypeptide is "substantially pure," "substantially homogeneous" or "substantially purified" when at least about 60% to 75% of a sample exhibits a single species of polypeptide. The polypeptide or protein may be monomeric or multimeric. A substantially pure polypeptide or protein will typically comprise about 50%, 60%, 70%, 80% or 90% W/W of a protein sample, more usually about 95%, and preferably will be over 99% pure. Protein purity or homogeneity may be determined by a number of means well known in the art, such as polyacrylamide gel electrophoresis of a protein sample, followed by visualizing a single polypeptide band upon staining the gel with a stain well known in the art. For certain purposes, higher resolution may be provided by using HPLC or other means well known in the art for purification.

The term "fragment" when used herein with respect to polypeptides of the present invention refers to a polypeptide that has an amino-terminal and/or carboxy-terminal deletion compared to a full-length CSP. In a preferred embodiment, the fragment is a contiguous sequence in which the amino acid sequence of the fragment is identical to the corresponding positions in the naturally occurring polypeptide. Fragments typically are at least 5, 6, 7, 8, 9 or 10 amino acids long, preferably at least 12, 14, 16 or 18 amino acids long, more preferably at least 20 amino acids long, more preferably at least 25, 30, 35, 40 or 45, amino acids, even more preferably at least 50 or 60 amino acids long, and even more preferably at least 70 amino acids long.

A "derivative" when used herein with respect to polypeptides of the present invention refers to a polypeptide which is substantially similar in primary structural sequence to a CSP but which includes, *e.g.*, *in vivo* or *in vitro* chemical and biochemical modifications that are not found in the CSP. Such modifications include, for example, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. Other modifications include, *e.g.*, labeling with radionuclides, and various enzymatic modifications, as will be readily appreciated by those skilled in the art. A variety of methods for labeling polypeptides and of substituents or labels useful for such purposes are well known in the art, and include radioactive isotopes such as ^{125}I , ^{32}P , ^{35}S , ^{14}C and ^3H , ligands which bind to labeled antiligands (*e.g.*, antibodies), fluorophores, chemiluminescent agents, enzymes, and antiligands which can serve as specific binding pair members for a labeled ligand. The choice of label depends on the sensitivity required, ease of conjugation with the primer, stability requirements, and available instrumentation. Methods for labeling polypeptides are well known in the art. *See* Ausubel (1992), *supra*; Ausubel (1999), *supra*.

The term "fusion protein" refers to polypeptides of the present invention coupled to a heterologous amino acid sequence. Fusion proteins are useful because they can be constructed to contain two or more desired functional elements from two or more different proteins. A fusion protein comprises at least 10 contiguous amino acids from a polypeptide of interest, more preferably at least 20 or 30 amino acids, even more preferably at least 40, 50 or 60 amino acids, yet more preferably at least 75, 100 or 125 amino acids. Fusion proteins can be produced recombinantly by constructing a nucleic acid sequence that encodes the polypeptide or a fragment thereof in frame with a nucleic acid sequence encoding a different protein or peptide and then expressing the fusion protein. Alternatively, a fusion protein can be produced chemically by crosslinking the polypeptide or a fragment thereof to another protein.

The term "analog" refers to both polypeptide analogs and non-peptide analogs. The term "polypeptide analog" as used herein refers to a polypeptide that is comprised of a segment of at least 25 amino acids that has substantial identity to a portion of an amino acid sequence but which contains non-natural amino acids or non-natural inter-residue bonds. In a preferred embodiment, the analog has the same or similar biological activity as the native polypeptide. Typically, polypeptide analogs comprise a conservative amino acid substitution (or insertion or deletion) with respect to the naturally occurring sequence. Analogs typically are at least 20 amino acids long, preferably at least 50 amino acids long or longer, and can often be as long as a full-length naturally occurring polypeptide.

The term "non-peptide analog" refers to a compound with properties that are analogous to those of a reference polypeptide. A non-peptide compound may also be termed a "peptide mimetic" or a "peptidomimetic." Such compounds are often developed with the aid of computerized molecular modeling. Peptide mimetics that are structurally similar to useful peptides may be used to produce an equivalent effect. Generally, peptidomimetics are structurally similar to a paradigm polypeptide (*i.e.*, a polypeptide that has a desired biochemical property or pharmacological activity), but have one or more peptide linkages optionally replaced by a linkage selected from the group consisting of: --CH₂NH--, --CH₂S--, --CH₂-CH₂--, --CH=CH--(cis and trans), --COCH₂--, --CH(OH)CH₂--, and --CH₂SO--, by methods well known in the art. Systematic substitution of one or more amino acids of a consensus sequence with a D-amino acid of the same type (*e.g.*, D-lysine in place of L-lysine) may also be used to generate more

stable peptides. In addition, constrained peptides comprising a consensus sequence or a substantially identical consensus sequence variation may be generated by methods known in the art (Rizo *et al.*, *Ann. Rev. Biochem.* 61:387-418 (1992)). For example, one may add internal cysteine residues capable of forming intramolecular disulfide bridges which
5 cyclize the peptide.

The term "mutant" or "mutein" when referring to a polypeptide of the present invention relates to an amino acid sequence containing substitutions, insertions or deletions of one or more amino acids compared to the amino acid sequence of a CSP. A mutein may have one or more amino acid point substitutions, in which a single amino acid
10 at a position has been changed to another amino acid, one or more insertions and/or deletions, in which one or more amino acids are inserted or deleted, respectively, in the sequence of the naturally occurring protein, and/or truncations of the amino acid sequence at either or both the amino or carboxy termini. Further, a mutein may have the same or different biological activity as the naturally occurring protein. For instance, a mutein may
15 have an increased or decreased biological activity. A mutein has at least 50% sequence similarity to the wild type protein, preferred is 60% sequence similarity, more preferred is 70% sequence similarity. Even more preferred are muteins having 80%, 85% or 90% sequence similarity to a CSP. In an even more preferred embodiment, a mutein exhibits 95% sequence identity, even more preferably 97%, even more preferably 98% and even
20 more preferably 99%. Sequence similarity may be measured by any common sequence analysis algorithm, such as GAP or BESTFIT or other variation Smith-Waterman alignment. *See*, T. F. Smith and M. S. Waterman, *J. Mol. Biol.* 147:195-197 (1981) and W.R. Pearson, *Genomics* 11:635-650 (1991).

Preferred amino acid substitutions are those which: (1) reduce susceptibility to
25 proteolysis, (2) reduce susceptibility to oxidation, (3) alter binding affinity for forming protein complexes, (4) alter binding affinity or enzymatic activity, and (5) confer or modify other physicochemical or functional properties of such analogs. For example, single or multiple amino acid substitutions (preferably conservative amino acid substitutions) may be made in the naturally occurring sequence (preferably in the portion
30 of the polypeptide outside the domain(s) forming intermolecular contacts. In a preferred embodiment, the amino acid substitutions are moderately conservative substitutions or conservative substitutions. In a more preferred embodiment, the amino acid substitutions are conservative substitutions. A conservative amino acid substitution should not

substantially change the structural characteristics of the parent sequence (*e.g.*, a replacement amino acid should not tend to disrupt a helix that occurs in the parent sequence, or disrupt other types of secondary structure that characterize the parent sequence). Examples of art-recognized polypeptide secondary and tertiary structures are described in Creighton (ed.), Proteins, Structures and Molecular Principles, W. H. Freeman and Company (1984); Branden *et al.* (ed.), Introduction to Protein Structure, Garland Publishing (1991); Thornton *et al.*, *Nature* 354:105-106 (1991).

As used herein, the twenty conventional amino acids and their abbreviations follow conventional usage. See Golub *et al.* (eds.), Immunology - A Synthesis 2nd Ed., Sinauer Associates (1991). Stereoisomers (*e.g.*, D-amino acids) of the twenty conventional amino acids, unnatural amino acids such as α -, α -disubstituted amino acids, N-alkyl amino acids, and other unconventional amino acids may also be suitable components for polypeptides of the present invention. Examples of unconventional amino acids include: 4-hydroxyproline, γ -carboxyglutamate, ϵ -N,N,N-trimethyllysine, ϵ -N-acetyllysine, O-phosphoserine, N-acetylserine, N-formylmethionine, 3-methylhistidine, 5-hydroxylysine, s-N-methylarginine, and other similar amino acids and imino acids (*e.g.*, 4-hydroxyproline). In the polypeptide notation used herein, the lefthand direction is the amino terminal direction and the right hand direction is the carboxy-terminal direction, in accordance with standard usage and convention.

By "homology" or "homologous" when referring to a polypeptide of the present invention it is meant polypeptides from different organisms with a similar sequence to the encoded amino acid sequence of a CSP and a similar biological activity or function. Although two polypeptides are said to be "homologous," this does not imply that there is necessarily an evolutionary relationship between the polypeptides. Instead, the term "homologous" is defined to mean that the two polypeptides have similar amino acid sequences and similar biological activities or functions. In a preferred embodiment, a homologous polypeptide is one that exhibits 50% sequence similarity to CSP, preferred is 60% sequence similarity, more preferred is 70% sequence similarity. Even more preferred are homologous polypeptides that exhibit 80%, 85% or 90% sequence similarity to a CSP. In yet a more preferred embodiment, a homologous polypeptide exhibits 95%, 97%, 98% or 99% sequence similarity.

When "sequence similarity" is used in reference to polypeptides, it is recognized that residue positions that are not identical often differ by conservative amino acid

substitutions. In a preferred embodiment, a polypeptide that has "sequence similarity" comprises conservative or moderately conservative amino acid substitutions. A "conservative amino acid substitution" is one in which an amino acid residue is substituted by another amino acid residue having a side chain (R group) with similar chemical properties (e.g., charge or hydrophobicity). In general, a conservative amino acid substitution will not substantially change the functional properties of a protein. In cases where two or more amino acid sequences differ from each other by conservative substitutions, the percent sequence identity or degree of similarity may be adjusted upwards to correct for the conservative nature of the substitution. Means for making this adjustment are well known to those of skill in the art. See, e.g., Pearson, *Methods Mol. Biol.* 24: 307-31 (1994).

For instance, the following six groups each contain amino acids that are conservative substitutions for one another:

- 1) Serine (S), Threonine (T);
- 2) Aspartic Acid (D), Glutamic Acid (E);
- 3) Asparagine (N), Glutamine (Q);
- 4) Arginine (R), Lysine (K);
- 5) Isoleucine (I), Leucine (L), Methionine (M), Alanine (A), Valine (V), and
- 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

Alternatively, a conservative replacement is any change having a positive value in the PAM250 log-likelihood matrix disclosed in Gonnet *et al.*, *Science* 256: 1443-45 (1992). A "moderately conservative" replacement is any change having a nonnegative value in the PAM250 log-likelihood matrix.

Sequence similarity for polypeptides, which is also referred to as sequence identity, is typically measured using sequence analysis software. Protein analysis software matches similar sequences using measures of similarity assigned to various substitutions, deletions and other modifications, including conservative amino acid substitutions. For instance, GCG contains programs such as "Gap" and "Bestfit" which can be used with default parameters to determine sequence homology or sequence identity between closely related polypeptides, such as homologous polypeptides from different species of organisms or between a wild type protein and a mutein thereof. See, e.g., GCG Version 6.1. Other programs include FASTA, discussed *supra*.

A preferred algorithm when comparing a sequence of the invention to a database containing a large number of sequences from different organisms is the computer program BLAST, especially blastp or tblastn. See, e.g., Altschul *et al.*, *J. Mol. Biol.* 215: 403-410 (1990); Altschul *et al.*, *Nucleic Acids Res.* 25:3389-402 (1997). Preferred parameters for
 5 blastp are:

	Expectation value:	10 (default)
	Filter:	seg (default)
	Cost to open a gap:	11 (default)
	Cost to extend a gap:	1 (default)
10	Max. alignments:	100 (default)
	Word size:	11 (default)
	No. of descriptions:	100 (default)
	Penalty Matrix:	BLOSUM62

The length of polypeptide sequences compared for homology will generally be at
 15 least about 16 amino acid residues, usually at least about 20 residues, more usually at least about 24 residues, typically at least about 28 residues, and preferably more than about 35 residues. When searching a database containing sequences from a large number of different organisms, it is preferable to compare amino acid sequences.

Algorithms other than blastp for database searching using amino acid sequences
 20 are known in the art. For instance, polypeptide sequences can be compared using FASTA, a program in GCG Version 6.1. FASTA (e.g., FASTA2 and FASTA3) provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences (Pearson (1990), *supra*; Pearson (2000), *supra*. For example, percent sequence identity between amino acid sequences can be determined using FASTA
 25 with its default or recommended parameters (a word size of 2 and the PAM250 scoring matrix), as provided in GCG Version 6.1.

An "antibody" refers to an intact immunoglobulin, or to an antigen-binding portion thereof that competes with the intact antibody for specific binding to a molecular species, e.g., a polypeptide of the instant invention. Antigen-binding portions may be produced by
 30 recombinant DNA techniques or by enzymatic or chemical cleavage of intact antibodies. Antigen-binding portions include, *inter alia*, Fab, Fab', F(ab')₂, Fv, dAb, and complementarity determining region (CDR) fragments, single-chain antibodies (scFv), chimeric antibodies, diabodies and polypeptides that contain at least a portion of an

immunoglobulin that is sufficient to confer specific antigen binding to the polypeptide. A Fab fragment is a monovalent fragment consisting of the VL, VH, CL and CH1 domains; a F(ab')₂ fragment is a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; a Fd fragment consists of the VH and CH1 domains; a Fv fragment consists of the VL and VH domains of a single arm of an antibody; and a dAb fragment consists of a VH domain. See, e.g., Ward *et al.*, *Nature* 341: 544-546 (1989).

By "bind specifically" and "specific binding" as used herein it is meant the ability of the antibody to bind to a first molecular species in preference to binding to other molecular species with which the antibody and first molecular species are admixed. An antibody is said to "recognize" a first molecular species when it can bind specifically to that first molecular species.

A single-chain antibody (scFv) is an antibody in which VL and VH regions are paired to form a monovalent molecule via a synthetic linker that enables them to be made as a single protein chain. See, e.g., Bird *et al.*, *Science* 242: 423-426 (1988); Huston *et al.*, *Proc. Natl. Acad. Sci. USA* 85: 5879-5883 (1988). Diabodies are bivalent, bispecific antibodies in which VH and VL domains are expressed on a single polypeptide chain, but using a linker that is too short to allow for pairing between the two domains on the same chain, thereby forcing the domains to pair with complementary domains of another chain and creating two antigen binding sites. See e.g., Holliger *et al.*, *Proc. Natl. Acad. Sci. USA* 90: 6444-6448 (1993); Poljak *et al.*, *Structure* 2: 1121-1123 (1994). One or more CDRs may be incorporated into a molecule either covalently or noncovalently to make it an immunoadhesin. An immunoadhesin may incorporate the CDR(s) as part of a larger polypeptide chain, may covalently link the CDR(s) to another polypeptide chain, or may incorporate the CDR(s) noncovalently. The CDRs permit the immunoadhesin to specifically bind to a particular antigen of interest. A chimeric antibody is an antibody that contains one or more regions from one antibody and one or more regions from one or more other antibodies.

An antibody may have one or more binding sites. If there is more than one binding site, the binding sites may be identical to one another or may be different. For instance, a naturally occurring immunoglobulin has two identical binding sites, a single-chain antibody or Fab fragment has one binding site, while a "bispecific" or "bifunctional" antibody has two different binding sites.

An "isolated antibody" is an antibody that (1) is not associated with naturally-associated components, including other naturally-associated antibodies, that accompany it in its native state, (2) is free of other proteins from the same species, (3) is expressed by a cell from a different species, or (4) does not occur in nature. It is known that purified
5 proteins, including purified antibodies, may be stabilized with non-naturally-associated components. The non-naturally-associated component may be a protein, such as albumin (e.g., BSA) or a chemical such as polyethylene glycol (PEG).

A "neutralizing antibody" or "an inhibitory antibody" is an antibody that inhibits the activity of a polypeptide or blocks the binding of a polypeptide to a ligand that
10 normally binds to it. An "activating antibody" is an antibody that increases the activity of a polypeptide.

The term "epitope" includes any protein determinant capable of specific binding to an immunoglobulin or T-cell receptor. Epitopic determinants usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains
15 and usually have specific three-dimensional structural characteristics, as well as specific charge characteristics. An antibody is said to specifically bind an antigen when the dissociation constant is less than $1\ \mu\text{M}$, preferably less than $100\ \text{nM}$ and most preferably less than $10\ \text{nM}$.

The term "patient" includes human and veterinary subjects.

20 Throughout this specification and claims, the word "comprise," or variations such as "comprises" or "comprising," will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers.

The term "colon specific" refers to a nucleic acid molecule or polypeptide that is expressed predominantly in the colon as compared to other tissues in the body. In a
25 preferred embodiment, a "colon specific" nucleic acid molecule or polypeptide is detected at a level that is 1.5-fold higher than any other tissue in the body. In a more preferred embodiment, the "colon specific" nucleic acid molecule or polypeptide is detected at a level that is 2-fold higher than any other tissue in the body, more preferably 5-fold higher, still more preferably at least 10-fold, 15-fold, 20-fold, 25-fold, 50-fold or 100-fold higher
30 than any other tissue in the body. Nucleic acid molecule levels may be measured by nucleic acid hybridization, such as Northern blot hybridization, or quantitative PCR. Polypeptide levels may be measured by any method known to accurately quantitate protein levels, such as Western blot analysis.

Nucleic Acid Molecules, Regulatory Sequences, Vectors, Host Cells and Recombinant
Methods of Making Polypeptides

Nucleic Acid Molecules

One aspect of the invention provides isolated nucleic acid molecules that are
5 specific to the colon or to colon cells or tissue or that are derived from such nucleic acid
molecules. These isolated colon specific nucleic acids (CSNAs) may comprise cDNA
genomic DNA, RNA, or a combination thereof, a fragment of one of these nucleic acids,
or may be a non-naturally occurring nucleic acid molecule. A CSNA may be derived from
an animal. In a preferred embodiment, the CSNA is derived from a human or other
10 mammal. In a more preferred embodiment, the CSNA is derived from a human or other
primate. In an even more preferred embodiment, the CSNA is derived from a human.

In a preferred embodiment, the nucleic acid molecule encodes a polypeptide that
is specific to colon, a colon-specific polypeptide (CSP). In a more preferred embodiment,
the nucleic acid molecule encodes a polypeptide that comprises an amino acid sequence of
15 SEQ ID NO: 96-237. In another highly preferred embodiment, the nucleic acid molecule
comprises a nucleic acid sequence of SEQ ID NO: 1-95. Nucleotide sequences of the
instantly-described nucleic acid molecules were determined by assembling several DNA
molecules from either public or proprietary databases. Some of the underlying DNA
sequences are the result, directly or indirectly, of at least one enzymatic polymerization
20 reaction (e.g., reverse transcription and/or polymerase chain reaction) using an automated
sequencer (such as the MegaBACE™ 1000, Amersham Biosciences, Sunnyvale, CA,
USA).

Nucleic acid molecules of the present invention may also comprise sequences that
selectively hybridize to a nucleic acid molecule encoding a CSNA or a complement or
25 antisense thereof. The hybridizing nucleic acid molecule may or may not encode a
polypeptide or may or may not encode a CSP. However, in a preferred embodiment, the
hybridizing nucleic acid molecule encodes a CSP. In a more preferred embodiment, the
invention provides a nucleic acid molecule that selectively hybridizes to a nucleic acid
molecule or the antisense sequence of a nucleic acid molecule that encodes a polypeptide
30 comprising an amino acid sequence of SEQ ID NO: 96-237. In an even more preferred
embodiment, the invention provides a nucleic acid molecule that selectively hybridizes to
a nucleic acid molecule comprising the nucleic acid sequence of SEQ ID NO: 1-95 or the
antisense sequence thereof. Preferably, the nucleic acid molecule selectively hybridizes to

a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a CSP under low stringency conditions. More preferably, the nucleic acid molecule selectively hybridizes to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a CSP under moderate stringency conditions. Most preferably, the nucleic acid molecule selectively hybridizes to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a CSP under high stringency conditions. In a preferred embodiment, the nucleic acid molecule hybridizes under low, moderate or high stringency conditions to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a polypeptide comprising an amino acid sequence of SEQ ID NO: 96-237. In a more preferred embodiment, the nucleic acid molecule hybridizes under low, moderate or high stringency conditions to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule comprising a nucleic acid sequence selected from SEQ ID NO: 1-95.

Nucleic acid molecules of the present invention may also comprise nucleic acid sequences that exhibit substantial sequence similarity to a nucleic acid encoding a CSP or a complement of the encoding nucleic acid molecule. In this embodiment, it is preferred that the nucleic acid molecule exhibit substantial sequence similarity to a nucleic acid molecule encoding human CSP. More preferred is a nucleic acid molecule exhibiting substantial sequence similarity to a nucleic acid molecule encoding a polypeptide having an amino acid sequence of SEQ ID NO: 96-237. By substantial sequence similarity it is meant a nucleic acid molecule having at least 60%, more preferably at least 70%, even more preferably at least 80% and even more preferably at least 85% sequence identity with a nucleic acid molecule encoding a CSP, such as a polypeptide having an amino acid sequence of SEQ ID NO: 96-237. In a more preferred embodiment, the similar nucleic acid molecule is one that has at least 90%, more preferably at least 95%, more preferably at least 97%, even more preferably at least 98%, and still more preferably at least 99% sequence identity with a nucleic acid molecule encoding a CSP. Most preferred in this embodiment is a nucleic acid molecule that has at least 99.5%, 99.6%, 99.7%, 99.8% or 99.9% sequence identity with a nucleic acid molecule encoding a CSP.

The nucleic acid molecules of the present invention are also inclusive of those exhibiting substantial sequence similarity to a CSNA or its complement. In this embodiment, it is preferred that the nucleic acid molecule exhibit substantial sequence similarity to a nucleic acid molecule having a nucleic acid sequence of SEQ ID NO: 1-95.

By substantial sequence similarity it is meant a nucleic acid molecule that has at least 60%, more preferably at least 70%, even more preferably at least 80% and even more preferably at least 85% sequence identity with a CSNA, such as one having a nucleic acid sequence of SEQ ID NO: 1-95. More preferred is a nucleic acid molecule that has at least
5 90%, more preferably at least 95%, more preferably at least 97%, even more preferably at least 98%, and still more preferably at least 99% sequence identity with a CSNA. Most preferred is a nucleic acid molecule that has at least 99.5%, 99.6%, 99.7%, 99.8% or 99.9% sequence identity with a CSNA.

Nucleic acid molecules that exhibit substantial sequence similarity are inclusive of
10 sequences that exhibit sequence identity over their entire length to a CSNA or to a nucleic acid molecule encoding a CSP, as well as sequences that are similar over only a part of its length. In this case, the part is at least 50 nucleotides of the CSNA or the nucleic acid molecule encoding a CSP, preferably at least 100 nucleotides, more preferably at least 150 or 200 nucleotides, even more preferably at least 250 or 300 nucleotides, still more
15 preferably at least 400 or 500 nucleotides.

The substantially similar nucleic acid molecule may be a naturally occurring one that is derived from another species, especially one derived from another primate, wherein the similar nucleic acid molecule encodes an amino acid sequence that exhibits significant sequence identity to that of SEQ ID NO: 96-237 or demonstrates significant sequence
20 identity to the nucleotide sequence of SEQ ID NO: 1-95. The similar nucleic acid molecule may also be a naturally occurring nucleic acid molecule from a human, when the CSNA is a member of a gene family. The similar nucleic acid molecule may also be a naturally occurring nucleic acid molecule derived from a non-primate, mammalian species, including without limitation, domesticated species, *e.g.*, dog, cat, mouse, rat,
25 rabbit, hamster, cow, horse and pig; and wild animals, *e.g.*, monkey, fox, lions, tigers, bears, giraffes, zebras, etc. The substantially similar nucleic acid molecule may also be a naturally occurring nucleic acid molecule derived from a non-mammalian species, such as birds or reptiles. The naturally occurring substantially similar nucleic acid molecule may be isolated directly from humans or other species. In another embodiment, the
30 substantially similar nucleic acid molecule may be one that is experimentally produced by random mutation of a nucleic acid molecule. In another embodiment, the substantially similar nucleic acid molecule may be one that is experimentally produced by directed

mutation of a CSNA. In a preferred embodiment, the substantially similar nucleic acid molecule is a CSNA.

The nucleic acid molecules of the present invention are also inclusive of allelic variants of a CSNA or a nucleic acid encoding a CSP. For example, single nucleotide polymorphisms (SNPs) occur frequently in eukaryotic genomes and the sequence
5 determined from one individual of a species may differ from other allelic forms present within the population. More than 1.4 million SNPs have already been identified in the human genome, International Human Genome Sequencing Consortium, *Nature* 409: 860-921 (2001) – Variants with small deletions and insertions of more than a single nucleotide
10 are also found in the general population, and often do not alter the function of the protein. In addition, amino acid substitutions occur frequently among natural allelic variants, and often do not substantially change protein function.

In a preferred embodiment, the allelic variant is a variant of a gene, wherein the gene is transcribed into a mRNA that encodes a CSP. In a more preferred embodiment,
15 the gene is transcribed into a mRNA that encodes a CSP comprising an amino acid sequence of SEQ ID NO: 96-237. In another preferred embodiment, the allelic variant is a variant of a gene, wherein the gene is transcribed into a mRNA that is a CSNA. In a more preferred embodiment, the gene is transcribed into an mRNA that comprises the nucleic acid sequence of SEQ ID NO: 1-95. Also preferred is that the allelic variant be a
20 naturally occurring allelic variant in the species of interest, particularly human.

Nucleic acid molecules of the present invention are also inclusive of nucleic acid sequences comprising a part of a nucleic acid sequence of the instant invention. The part may or may not encode a polypeptide, and may or may not encode a polypeptide that is a CSP. In a preferred embodiment, the part encodes a CSP. In one embodiment, the
25 nucleic acid molecule comprises a part of a CSNA. In another embodiment, the nucleic acid molecule comprises a part of a nucleic acid molecule that hybridizes or exhibits substantial sequence similarity to a CSNA. In another embodiment, the nucleic acid molecule comprises a part of a nucleic acid molecule that is an allelic variant of a CSNA. In yet another embodiment, the nucleic acid molecule comprises a part of a nucleic acid
30 molecule that encodes a CSP. A part comprises at least 10 nucleotides, more preferably at least 15, 17, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400 or 500 nucleotides. The maximum size of a nucleic acid part is one nucleotide shorter than the sequence of the nucleic acid molecule encoding the full-length protein.

Nucleic acid molecules of the present invention are also inclusive of nucleic acid sequences that encode fusion proteins, homologous proteins, polypeptide fragments, muteins and polypeptide analogs, as described *infra*.

5 Nucleic acid molecules of the present invention are also inclusive of nucleic acid sequences containing modifications of the native nucleic acid molecule. Examples of such modifications include, but are not limited to, nonnative internucleoside bonds, post-synthetic modifications or altered nucleotide analogues. One having ordinary skill in the art would recognize that the type of modification that may be made will depend upon the intended use of the nucleic acid molecule. For instance, when the nucleic acid molecule is
10 used as a hybridization probe, the range of such modifications will be limited to those that permit sequence-discriminating base pairing of the resulting nucleic acid. When used to direct expression of RNA or protein *in vitro* or *in vivo*, the range of such modifications will be limited to those that permit the nucleic acid to function properly as a polymerization substrate. When the isolated nucleic acid is used as a therapeutic agent,
15 the modifications will be limited to those that do not confer toxicity upon the isolated nucleic acid.

Accordingly, in one embodiment, a nucleic acid molecule may include nucleotide analogues that incorporate labels that are directly detectable, such as radiolabels or fluorophores, or nucleotide analogues that incorporate labels that can be visualized in a
20 subsequent reaction, such as biotin or various haptens. The labeled nucleic acid molecules are particularly useful as hybridization probes.

Common radiolabeled analogues include those labeled with ^{33}P , ^{32}P , and ^{35}S , such as α - ^{32}P -dATP, α - ^{32}P -dCTP, α - ^{32}P -dGTP, α - ^{32}P -dTTP, α - ^{32}P -3'dATP, α - ^{32}P -ATP, α - ^{32}P -CTP, α - ^{32}P -GTP, α - ^{32}P -UTP, α - ^{35}S -dATP, γ - ^{35}S -GTP, γ - ^{33}P -dATP, and the like.

25 Commercially available fluorescent nucleotide analogues readily incorporated into the nucleic acids of the present invention include Cy3-dCTP, Cy3-dUTP, Cy5-dCTP, Cy3-dUTP (Amersham Biosciences, Piscataway, New Jersey, USA), fluorescein-12-dUTP, tetramethylrhodamine-6-dUTP, Texas Red®-5-dUTP, Cascade Blue®-7-dUTP, BODIPY® FL-14-dUTP, BODIPY® TMR-14-dUTP, BODIPY® TR-14-dUTP,
30 Rhodamine Green™-5-dUTP, Oregon Green® 488-5-dUTP, Texas Red®-12-dUTP, BODIPY® 630/650-14-dUTP, BODIPY® 650/665-14-dUTP, Alexa Fluor® 488-5-dUTP, Alexa Fluor® 532-5-dUTP, Alexa Fluor® 568-5-dUTP, Alexa Fluor® 594-5-dUTP, Alexa Fluor® 546-14-dUTP, fluorescein-12-UTP, tetramethylrhodamine-6-UTP, Texas

Red®-5-UTP, Cascade Blue®-7-UTP, BODIPY® FL-14-UTP, BODIPY® TMR-14-UTP, BODIPY® TR-14-UTP, Rhodamine Green™-5-UTP, Alexa Fluor® 488-5-UTP, Alexa Fluor® 546-14-UTP (Molecular Probes, Inc. Eugene, OR, USA). One may also custom synthesize nucleotides having other fluorophores. See Henegariu *et al.*, *Nature*

5 *Biotechnol.* 18: 345-348 (2000).

Haptens that are commonly conjugated to nucleotides for subsequent labeling include biotin (biotin-11-dUTP, Molecular Probes, Inc., Eugene, OR, USA; biotin-21-UTP, biotin-21-dUTP, Clontech Laboratories, Inc., Palo Alto, CA, USA), digoxigenin (DIG-11-dUTP, alkali labile, DIG-11-UTP, Roche Diagnostics Corp., Indianapolis, IN, USA), and dinitrophenyl (dinitrophenyl-11-dUTP, Molecular Probes, Inc., Eugene, OR, USA).

Nucleic acid molecules of the present invention can be labeled by incorporation of labeled nucleotide analogues into the nucleic acid. Such analogues can be incorporated by enzymatic polymerization, such as by nick translation, random priming, polymerase chain reaction (PCR), terminal transferase tailing, and end-filling of overhangs, for DNA molecules, and *in vitro* transcription driven, *e.g.*, from phage promoters, such as T7, T3, and SP6, for RNA molecules. Commercial kits are readily available for each such labeling approach. Analogues can also be incorporated during automated solid phase chemical synthesis. Labels can also be incorporated after nucleic acid synthesis, with the 5' phosphate and 3' hydroxyl providing convenient sites for post-synthetic covalent attachment of detectable labels.

Other post-synthetic approaches also permit internal labeling of nucleic acids. For example, fluorophores can be attached using a cisplatin reagent that reacts with the N7 of guanine residues (and, to a lesser extent, adenine bases) in DNA, RNA, and Peptide Nucleic Acids (PNA) to provide a stable coordination complex between the nucleic acid and fluorophore label (Universal Linkage System) (available from Molecular Probes, Inc., Eugene, OR, USA and Amersham Pharmacia Biotech, Piscataway, NJ, USA); see Alers *et al.*, *Genes, Chromosomes & Cancer* 25: 301- 305 (1999); Jelsma *et al.*, *J. NIH Res.* 5: 82 (1994); Van Belkum *et al.*, *BioTechniques* 16: 148-153 (1994). Alternatively, nucleic acids can be labeled using a disulfide-containing linker (FastTag™ Reagent, Vector Laboratories, Inc., Burlingame, CA, USA) that is photo- or thermally coupled to the target nucleic acid using aryl azide chemistry; after reduction, a free thiol is available for coupling to a hapten, fluorophore, sugar, affinity ligand, or other marker.

One or more independent or interacting labels can be incorporated into the nucleic acid molecules of the present invention. For example, both a fluorophore and a moiety that in proximity thereto acts to quench fluorescence can be included to report specific hybridization through release of fluorescence quenching or to report exonucleotidic excision. See, e.g., Tyagi *et al.*, *Nature Biotechnol.* 14: 303-308 (1996); Tyagi *et al.*, *Nature Biotechnol.* 16: 49-53 (1998); Sokol *et al.*, *Proc. Natl. Acad. Sci. USA* 95: 11538-11543 (1998); Kostrikis *et al.*, *Science* 279: 1228-1229 (1998); Marras *et al.*, *Genet. Anal.* 14: 151-156 (1999); Holland *et al.*, *Proc. Natl. Acad. Sci. USA* 88: 7276-7280 (1991); Heid *et al.*, *Genome Res.* 6(10): 986-94 (1996); Kuimelis *et al.*, *Nucleic Acids Symp. Ser.* (37): 255-6 (1997); and U.S. Patent Nos. 5,846,726, 5,925,517, 5,925,517, 5,723,591 and 5,538,848, the disclosures of which are incorporated herein by reference in their entireties.

Nucleic acid molecules of the present invention may also be modified by altering one or more native phosphodiester internucleoside bonds to more nuclease-resistant, internucleoside bonds. See Hartmann *et al.* (eds.), Manual of Antisense Methodology: Perspectives in Antisense Science, Kluwer Law International (1999); Stein *et al.* (eds.), Applied Antisense Oligonucleotide Technology, Wiley-Liss (1998); Chadwick *et al.* (eds.), Oligonucleotides as Therapeutic Agents – Symposium No. 209, John Wiley & Son Ltd (1997). Such altered internucleoside bonds are often desired for techniques or for targeted gene correction, Gamper *et al.*, *Nucl. Acids Res.* 28(21): 4332-4339 (2000). For double-stranded RNA inhibition which may utilize either natural ds RNA or ds RNA modified in its, sugar, phosphate or base, see Hannon, *Nature* 418(11): 244-251 (2002); Fire *et al.* in WO 99/32619; Tuschl *et al.* in US2002/0086356; Kruetzer *et al.* in WO 00/44895, the disclosures of which are incorporated herein by reference in their entirety. For circular antisense, see Kool in U.S. Patent No. 5,426,180, the disclosure of which is incorporated herein by reference in its entirety.

Modified oligonucleotide backbones include, without limitation, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, and boranophosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity

wherein the adjacent pairs of nucleoside units are linked 3'-5' to 5'-3' or 2'-5' to 5'-2'.

Representative U.S. Patents that teach the preparation of the above phosphorus-containing linkages include, but are not limited to, U.S. Patent Nos. 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; and 5,625,050, the disclosures of which are incorporated herein by reference in their entireties. In a preferred embodiment, the modified internucleoside linkages may be used for antisense techniques.

Other modified oligonucleotide backbones do not include a phosphorus atom, but have backbones that are formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom and alkyl or cycloalkyl internucleoside linkages, or one or more short chain heteroatomic or heterocyclic internucleoside linkages. These include those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfoxide and sulfone backbones; formacetyl and thioformacetyl backbones; methylene formacetyl and thioformacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH₂ component parts. Representative U.S. patents that teach the preparation of the above backbones include, but are not limited to, U.S. Patent Nos. 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437 and 5,677,439; the disclosures of which are incorporated herein by reference in their entireties.

In other preferred nucleic acid molecules, both the sugar and the internucleoside linkage are replaced with novel groups, such as peptide nucleic acids (PNA). In PNA compounds, the phosphodiester backbone of the nucleic acid is replaced with an amide-containing backbone, in particular by repeating N-(2-aminoethyl) glycine units linked by amide bonds. Nucleobases are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone, typically by methylene carbonyl linkages. PNA can be synthesized using a modified peptide synthesis protocol. PNA oligomers can be synthesized by both Fmoc and tBoc methods. Representative U.S. patents that teach the preparation of PNA compounds include, but are not limited to, U.S. Patent Nos.

5,539,082; 5,714,331; and 5,719,262, each of which is herein incorporated by reference in its entirety. Automated PNA synthesis is readily achievable on commercial synthesizers (see, e.g., "PNA User's Guide," Rev. 2, February 1998, Perseptive Biosystems Part No. 60138, Applied Biosystems, Inc., Foster City, CA). PNA molecules are advantageous for a number of reasons. First, because the PNA backbone is uncharged, PNA/DNA and PNA/RNA duplexes have a higher thermal stability than is found in DNA/DNA and DNA/RNA duplexes. The T_m of a PNA/DNA or PNA/RNA duplex is generally 1°C higher per base pair than the T_m of the corresponding DNA/DNA or DNA/RNA duplex (in 100 mM NaCl). Second, PNA molecules can also form stable PNA/DNA complexes at low ionic strength, under conditions in which DNA/DNA duplex formation does not occur. Third, PNA also demonstrates greater specificity in binding to complementary DNA because a PNA/DNA mismatch is more destabilizing than DNA/DNA mismatch. A single mismatch in mixed a PNA/DNA 15-mer lowers the T_m by 8–20°C (15°C on average). In the corresponding DNA/DNA duplexes, a single mismatch lowers the T_m by 4–16°C (11°C on average). Because PNA probes can be significantly shorter than DNA probes, their specificity is greater. Fourth, PNA oligomers are resistant to degradation by enzymes, and the lifetime of these compounds is extended both *in vivo* and *in vitro* because nucleases and proteases do not recognize the PNA polyamide backbone with nucleobase sidechains. See, e.g., Ray *et al.*, *FASEB J.* 14(9): 1041-60 (2000); Nielsen *et al.*, *Pharmacol Toxicol.* 86(1): 3-7 (2000); Larsen *et al.*, *Biochim Biophys Acta.* 1489(1): 159-66 (1999); Nielsen, *Curr. Opin. Struct. Biol.* 9(3): 353-7 (1999), and Nielsen, *Curr. Opin. Biotechnol.* 10(1): 71-5 (1999).

Nucleic acid molecules may be modified compared to their native structure throughout the length of the nucleic acid molecule or can be localized to discrete portions thereof. As an example of the latter, chimeric nucleic acids can be synthesized that have discrete DNA and RNA domains and that can be used for targeted gene repair and modified PCR reactions, as further described in, Misra *et al.*, *Biochem.* 37: 1917-1925 (1998); and Finn *et al.*, *Nucl. Acids Res.* 24: 3357-3363 (1996), and U.S. Patent Nos. 5,760,012 and 5,731,181, the disclosures of which are incorporated herein by reference in their entireties.

Unless otherwise specified, nucleic acid molecules of the present invention can include any topological conformation appropriate to the desired use; the term thus explicitly comprehends, among others, single-stranded, double-stranded, triplexed,

quadruplexed, partially double-stranded, partially-triplexed, partially-quadruplexed, branched, hairpinned, circular, and padlocked conformations. Padlocked conformations and their utilities are further described in Banér *et al.*, *Curr. Opin. Biotechnol.* 12: 11-15 (2001); Escude *et al.*, *Proc. Natl. Acad. Sci. USA* 14: 96(19):10603-7 (1999); and Nilsson
 5 *et al.*, *Science* 265(5181): 2085-8 (1994). Triplexed and quadruplexed conformations, and their utilities, are reviewed in Praseuth *et al.*, *Biochim. Biophys. Acta.* 1489(1): 181-206 (1999); Fox, *Curr. Med. Chem.* 7(1): 17-37 (2000); Kochetkova *et al.*, *Methods Mol. Biol.* 130: 189-201 (2000); Chan *et al.*, *J. Mol. Med.* 75(4): 267-82 (1997); Rowley *et al.*, *Mol Med* 5(10): 693-700 (1999); Kool, *Annu Rev Biophys Biomol Struct.* 25: 1-28 (1996).

10 *SNP Polymorphisms*

Commonly, sequence differences between individuals involve differences in single nucleotide positions. SNPs may account for 90% of human DNA polymorphism. Collins
et al., 8 *Genome Res.* 1229-31 (1998). SNPs include single base pair positions in genomic DNA at which different sequence alternatives (alleles) exist in a population. In addition,
 15 the least frequent allele generally must occur at a frequency of 1% or greater. DNA sequence variants with a reasonably high population frequency are observed approximately every 1,000 nucleotide across the genome, with estimates as high as 1 SNP per 350 base pairs. Wang *et al.*, 280 *Science* 1077-82 (1998); Harding *et al.*, 60 *Am. J. Human Genet.* 772-89 (1997); Taillon-Miller *et al.*, 8 *Genome Res.* 748-54 (1998); Cargill
 20 *et al.*, 22 *Nat. Genet.* 231-38 (1999); and Semple *et al.*, 16 *Bioinform. Disc. Note* 735-38 (2000). The frequency of SNPs varies with the type and location of the change. In base substitutions, two-thirds of the substitutions involve the C-T and G-A type. This variation in frequency can be related to 5-methylcytosine deamination reactions that occur frequently, particularly at CpG dinucleotides. Regarding location, SNPs occur at a much
 25 higher frequency in non-coding regions than in coding regions. Information on over one million variable sequences is already publicly available via the Internet and more such markers are available from commercial providers of genetic information. Kwok and Gu, 5 *Med. Today* 538-53 (1999).

Several definitions of SNPs exist. See, e.g., Brooks, 235 *Gene* 177-86 (1999). As
 30 used herein, the term "single nucleotide polymorphism" or "SNP" includes all single base variants, thus including nucleotide insertions and deletions in addition to single nucleotide substitutions. There are two types of nucleotide substitutions. A transition is the

replacement of one purine by another purine or one pyrimidine by another pyrimidine. A transversion is the replacement of a purine for a pyrimidine, or vice versa.

Numerous methods exist for detecting SNPs within a nucleotide sequence. A review of many of these methods can be found in Landegren *et al.*, 8 *Genome Res.* 769-76
5 (1998). For example, a SNP in a genomic sample can be detected by preparing a Reduced Complexity Genome (RCG) from the genomic sample, then analyzing the RCG for the presence or absence of a SNP. See, e.g., WO 00/18960 which is herein incorporated by reference in its entirety. Multiple SNPs in a population of target polynucleotides in parallel can be detected using, for example, the methods of WO 00/50869 which is herein
10 incorporated by reference in its entirety. Other SNP detection methods include the methods of U.S. Pat. Nos. 6,297,018 and 6,322,980 which are herein incorporated by reference in their entirety. Furthermore, SNPs can be detected by restriction fragment length polymorphism (RFLP) analysis. See, e.g., U.S. Pat. Nos. 5,324,631; 5,645,995 which are herein incorporated by reference in their entirety. RFLP analysis of SNPs,
15 however, is limited to cases where the SNP either creates or destroys a restriction enzyme cleavage site. SNPs can also be detected by direct sequencing of the nucleotide sequence of interest. In addition, numerous assays based on hybridization have also been developed to detect SNPs and mismatch distinction by polymerases and ligases. Several web sites provide information about SNPs including Ensembl on the World Wide Web at
20 ensemble.org, Sanger Institute on the World Wide Web at sanger.ac.uk/genetics/exon/, National Center for Biotechnology Information (NCBI) on the World Wide Web at ncbi.nlm.nih.gov/SNP/, The SNP Consortium Ltd. on the World Wide Web at snp.cshl.org. The chromosomal locations for the compositions disclosed herein are provided below. In addition, one of ordinary skill in the art could use a BLAST against
25 the genome or any of the databases cited above to find the chromosomal location. Another a preferred method to find the genomic coordinates and associated SNPs would be to use the BLAT tool (genome.ucsc.edu, Kent et al. 2001, The Human Genome Browser at UCSC, Genome Research 996-1006 or Kent 2002 BLAT —The BLAST -Like Alignment Tool Genome Reseach, 1-9). All web sites above were accessed December 3,
30 2003.

RNA interference

RNA interference refers to the process of sequence-specific post transcriptional gene silencing in animals mediated by short interfering RNAs (siRNA). Fire *et al.*, 1998, *Nature*, 391, 806. The corresponding process in plants is commonly referred to as post transcriptional gene silencing or RNA silencing and is also referred to as quelling in fungi.

- 5 The process of post transcriptional gene silencing is thought to be an evolutionarily conserved cellular defense mechanism used to prevent the expression of foreign genes which is commonly shared by diverse flora and phyla. Fire *et al.*, 1999, *Trends Genet.*, 15, 358. Such protection from foreign gene expression may have evolved in response to the production of double-stranded RNAs (dsRNA) derived from viral infection or the
- 10 random integration of transposon elements into a host genome via a cellular response that specifically destroys homologous single-stranded RNA or viral genomic RNA. The presence of dsRNA in cells triggers the RNAi response through a mechanism that has yet to be fully characterized. This mechanism appears to be different from the interferon response that results from dsRNA mediated activation of protein kinase PKR and 2',5'-
- 15 oligoadenylate synthetase resulting in non-specific cleavage of mRNA by ribonuclease L.

- The presence of long dsRNAs in cells stimulates the activity of a ribonuclease III enzyme referred to as dicer. Dicer is involved in the processing of the dsRNA into short pieces of dsRNA known as short interfering RNAs (siRNA). Berstein *et al.*, 2001, *Nature*, 409, 363. Short interfering RNAs derived from dicer activity are typically about
- 20 21-23 nucleotides in length and comprise about 19 base pair duplexes. Dicer has also been implicated in the excision of 21 and 22 nucleotide small temporal RNAs (stRNA) from precursor RNA of conserved structure that are implicated in translational control. Hutvagner *et al.*, 2001, *Science*, 293, 834. The RNAi response also features an
- 25 endonuclease complex containing a siRNA, commonly referred to as an RNA-induced silencing complex (RISC), which mediates cleavage of single-stranded RNA having sequence complementary to the antisense strand of the siRNA duplex. Cleavage of the target RNA takes place in the middle of the region complementary to the antisense strand of the siRNA duplex. Elbashir *et al.*, 2001, *Genes Dev.*, 15, 188.

- Short interfering RNA mediated RNAi has been studied in a variety of systems.
- 30 Fire *et al.*, 1998, *Nature*, 391, 806, were the first to observe RNAi in *C. Elegans*. Wianny and Goetz, 1999, *Nature Cell Biol.*, 2, 70, describe RNAi mediated by dsRNA in mouse embryos. Hammond *et al.*, 2000, *Nature*, 404, 293, describe RNAi in *Drosophila* cells

transfected with dsRNA. Elbashir *et al.*, 2001, *Nature*, 411, 494, describe RNAi induced by introduction of duplexes of synthetic 21-nucleotide RNAs in cultured mammalian cells including human embryonic kidney and HeLa cells. Recent work in *Drosophila* embryonic lysates (Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877) has revealed certain requirements for
5 siRNA length, structure, chemical composition, and sequence that are essential to mediate efficient RNAi activity. These studies have shown that 21 nucleotide siRNA duplexes are most active when containing two nucleotide 3'-overhangs. Furthermore, complete substitution of one or both siRNA strands with 2'-deoxy (2'-H) or 2'-O-methyl nucleotides abolishes RNAi activity, whereas substitution of the 3'-terminal siRNA overhang
10 nucleotides with deoxy nucleotides (2'-H) was shown to be tolerated. Single mismatch sequences in the center of the siRNA duplex were also shown to abolish RNAi activity. In addition, these studies also indicate that the position of the cleavage site in the target RNA is defined by the 5'-end of the siRNA guide sequence rather than the 3'-end. Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877. Other studies have indicated that a 5'-phosphate on the
15 target-complementary strand of a siRNA duplex is required for siRNA activity and that ATP is utilized to maintain the 5'-phosphate moiety on the siRNA. Nykanen *et al.*, 2001, *Cell*, 107, 309.

Studies have shown that replacing the 3'-overhanging segments of a 21-mer siRNA duplex having 2 nucleotide 3' overhangs with deoxyribonucleotides does not have an
20 adverse effect on RNAi activity. Replacing up to 4 nucleotides on each end of the siRNA with deoxyribonucleotides has been reported to be well tolerated whereas complete substitution with deoxyribonucleotides results in no RNAi activity. Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877. In addition, Elbashir *et al.*, *supra*, also report that substitution of siRNA with 2'-O-methyl nucleotides completely abolishes RNAi activity. Li *et al.*, WO
25 00/44914, and Beach *et al.*, WO 01/68836 both suggest that siRNA "may include modifications to either the phosphate-sugar back bone or the nucleoside to include at least one of a nitrogen or sulfur heteroatom", however neither application teaches to what extent these modifications are tolerated in siRNA molecules nor provides any examples of such modified siRNA. Kreutzer and Limmer, Canadian Patent Application No. 2,359,180, also
30 describe certain chemical modifications for use in dsRNA constructs in order to counteract activation of double-stranded RNA-dependent protein kinase PKR, specifically 2'-amino or 2'-O-methyl nucleotides, and nucleotides containing a 2'-O or 4'-C methylene bridge. However, Kreutzer and Limmer similarly fail to show to what extent these modifications

are tolerated in siRNA molecules nor do they provide any examples of such modified siRNA.

Parrish et al., 2000, *Molecular Cell*, 6, 1977-1087, tested certain chemical modifications targeting the unc-22 gene in *C. elegans* using long (>25 nt) siRNA transcripts. The authors describe the introduction of thiophosphate residues into these
5 siRNA transcripts by incorporating thiophosphate nucleotide analogs with T7 and T3 RNA polymerase and observed that "RNAs with two [phosphorothioate] modified bases also had substantial decreases in effectiveness as RNAi triggers; [phosphorothioate] modification of more than two residues greatly destabilized the RNAs in vitro and we
10 were not able to assay interference activities." Parrish et al. at 1081. The authors also tested certain modifications at the 2'-position of the nucleotide sugar in the long siRNA transcripts and observed that substituting deoxynucleotides for ribonucleotides "produced a substantial decrease in interference activity", especially in the case of Uridine to Thymidine and/or Cytidine to deoxy-Cytidine substitutions. Parrish et al. In addition, the
15 authors tested certain base modifications, including substituting 4-thiouracil, 5-bromouracil, 5-iodouracil, 3-(aminoallyl)uracil for uracil, and inosine for guanosine in sense and antisense strands of the siRNA, and found that whereas 4-thiouracil and 5-bromouracil were all well tolerated, inosine "produced a substantial decrease in interference activity" when incorporated in either strand. Incorporation of 5-iodouracil and
20 3-(aminoallyl)uracil in the antisense strand resulted in substantial decrease in RNAi activity as well.

Beach et al., WO 01/68836, describes specific methods for attenuating gene expression using endogenously derived dsRNA. Tuschl et al., WO 01/75164, describes a *Drosophila* in vitro RNAi system and the use of specific siRNA molecules for certain
25 functional genomic and certain therapeutic applications; although Tuschl, 2001, *Chem. Biochem.*, 2, 239-245, doubts that RNAi can be used to cure genetic diseases or viral infection due "to the danger of activating interferon response". Li et al., WO 00/44914, describes the use of specific dsRNAs for use in attenuating the expression of certain target genes. Zernicka-Goetz et al., WO 01/36646, describes certain methods for inhibiting the
30 expression of particular genes in mammalian cells using certain dsRNA molecules. Fire et al., WO 99/32619, U.S. Patent No. 6,506,559, the contents of which are hereby incorporated by reference in their entirety, describes particular methods for introducing

certain dsRNA molecules into cells for use in inhibiting gene expression. Plaetinck et al., WO 00/01846, describes certain methods for identifying specific genes responsible for conferring a particular phenotype in a cell using specific dsRNA molecules. Mello et al., WO 01/29058, describes the identification of specific genes involved in dsRNA mediated RNAi. Deschamps Depaillette et al., International PCT Publication No. WO 99/07409, describes specific compositions consisting of particular dsRNA molecules combined with certain anti-viral agents. Driscoll et al., International PCT Publication No. WO 01/49844, describes specific DNA constructs for use in facilitating gene silencing in targeted organisms. Parrish et al., 2000, Molecular Cell, 6, 1977-1087, describes specific chemically modified siRNA constructs targeting the unc-22 gene of *C. elegans*. Tuschl et al., International PCT Publication No. WO 02/44321, describe certain synthetic siRNA constructs.

Methods for Using Nucleic Acid Molecules as Probes and Primers

The isolated nucleic acid molecules of the present invention can be used as hybridization probes to detect, characterize, and quantify hybridizing nucleic acids in, and isolate hybridizing nucleic acids from, both genomic and transcript-derived nucleic acid samples. When free in solution, such probes are typically, but not invariably, detectably labeled; bound to a substrate, as in a microarray, such probes are typically, but not invariably unlabeled.

In one embodiment, the isolated nucleic acid molecules of the present invention can be used as probes to detect and characterize gross alterations in the gene of a CSNA, such as deletions, insertions, translocations, and duplications of the CSNA genomic locus through fluorescence *in situ* hybridization (FISH) to chromosome spreads. See, e.g., Andreeff *et al.* (eds.), Introduction to Fluorescence *In Situ* Hybridization: Principles and Clinical Applications, John Wiley & Sons (1999). The isolated nucleic acid molecules of the present invention can be used as probes to assess smaller genomic alterations using, e.g., Southern blot detection of restriction fragment length polymorphisms. The isolated nucleic acid molecules of the present invention can be used as probes to isolate genomic clones that include a nucleic acid molecule of the present invention, which thereafter can be restriction mapped and sequenced to identify deletions, insertions, translocations, and substitutions (single nucleotide polymorphisms, SNPs) at the sequence level.

Alternatively, detection techniques such as molecular beacons may be used, see Kostrikis *et al. Science* 279:1228-1229 (1998).

The isolated nucleic acid molecules of the present invention can also be used as probes to detect, characterize, and quantify CSNA in, and isolate CSNA from, transcript-derived nucleic acid samples. In one embodiment, the isolated nucleic acid molecules of the present invention can be used as hybridization probes to detect, characterize by length, and quantify mRNA by Northern blot of total or poly-A⁺-selected RNA samples. In another embodiment, the isolated nucleic acid molecules of the present invention can be used as hybridization probes to detect, characterize by location, and quantify mRNA by *in situ* hybridization to tissue sections. *See, e.g.,* Schwarczacher *et al., In Situ Hybridization*, Springer-Verlag New York (2000). In another preferred embodiment, the isolated nucleic acid molecules of the present invention can be used as hybridization probes to measure the representation of clones in a cDNA library or to isolate hybridizing nucleic acid molecules acids from cDNA libraries, permitting sequence level characterization of mRNAs that hybridize to CSNAs, including, without limitations, identification of deletions, insertions, substitutions, truncations, alternatively spliced forms and single nucleotide polymorphisms. In yet another preferred embodiment, the nucleic acid molecules of the instant invention may be used in microarrays.

All of the aforementioned probe techniques are well within the skill in the art, and are described at greater length in standard texts such as Sambrook (2001), *supra*; Ausubel (1999), *supra*; and Walker *et al. (eds.)*, The Nucleic Acids Protocols Handbook, Humana Press (2000).

In another embodiment, a nucleic acid molecule of the invention may be used as a probe or primer to identify and/or amplify a second nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of the invention. In this embodiment, it is preferred that the probe or primer be derived from a nucleic acid molecule encoding a CSP. More preferably, the probe or primer is derived from a nucleic acid molecule encoding a polypeptide having an amino acid sequence of SEQ ID NO: 96-237. Also preferred are probes or primers derived from a CSNA. More preferred are probes or primers derived from a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-95.

In general, a probe or primer is at least 10 nucleotides in length, more preferably at least 12, more preferably at least 14 and even more preferably at least 16 or 17 nucleotides

in length. In an even more preferred embodiment, the probe or primer is at least 18 nucleotides in length, even more preferably at least 20 nucleotides and even more preferably at least 22 nucleotides in length. Primers and probes may also be longer in length. For instance, a probe or primer may be 25 nucleotides in length, or may be 30, 40
5 or 50 nucleotides in length. Methods of performing nucleic acid hybridization using oligonucleotide probes are well known in the art. *See, e.g.,* Sambrook *et al.*, 1989, *supra*, Chapter 11 and pp. 11.31-11.32 and 11.40-11.44, which describes radiolabeling of short probes, and pp. 11.45-11.53, which describe hybridization conditions for oligonucleotide probes, including specific conditions for probe hybridization (pp. 11.50-11.51).

10 Methods of performing primer-directed amplification are also well known in the art. Methods for performing the polymerase chain reaction (PCR) are compiled, *inter alia*, in McPherson, PCR Basics: From Background to Bench, Springer Verlag (2000); Innis *et al.* (eds.), PCR Applications: Protocols for Functional Genomics, Academic Press (1999); Gelfand *et al.* (eds.), PCR Strategies, Academic Press (1998); Newton *et al.*, PCR,
15 Springer-Verlag New York (1997); Burke (ed.), PCR: Essential Techniques, John Wiley & Son Ltd (1996); White (ed.), PCR Cloning Protocols: From Molecular Cloning to Genetic Engineering, Vol. 67, Humana Press (1996); and McPherson *et al.* (eds.), PCR 2: A Practical Approach, Oxford University Press, Inc. (1995). Methods for performing RT-PCR are collected, *e.g.,* in Siebert *et al.* (eds.), Gene Cloning and Analysis by RT-PCR,
20 Eaton Publishing Company/Bio Techniques Books Division, 1998; and Siebert (ed.), PCR Technique: RT-PCR, Eaton Publishing Company/ BioTechniques Books (1995).

PCR and hybridization methods may be used to identify and/or isolate nucleic acid molecules of the present invention including allelic variants, homologous nucleic acid molecules and fragments. PCR and hybridization methods may also be used to identify,
25 amplify and/or isolate nucleic acid molecules of the present invention that encode homologous proteins, analogs, fusion proteins or muteins of the invention. Nucleic acid primers as described herein can be used to prime amplification of nucleic acid molecules of the invention, using transcript-derived or genomic DNA as the template.

These nucleic acid primers can also be used, for example, to prime single base
30 extension (SBE) for SNP detection (*See, e.g.,* U.S. Pat. No. 6,004,744, the disclosure of which is incorporated herein by reference in its entirety).

Isothermal amplification approaches, such as rolling circle amplification, are also now well-described. *See, e.g.,* Schweitzer *et al.*, *Curr. Opin. Biotechnol.* 12(1): 21-7

(2001); International Patent publications WO 97/19193 and WO 00/15779, and U.S. Patent Nos. 5,854,033 and 5,714,320, the disclosures of which are incorporated herein by reference in their entireties. Rolling circle amplification can be combined with other techniques to facilitate SNP detection. *See, e.g., Lizardi et al., Nature Genet.* 19(3):

5 225-32 (1998).

Nucleic acid molecules of the present invention may be bound to a substrate either covalently or noncovalently. The substrate can be porous or solid, planar or non-planar, unitary or distributed. The bound nucleic acid molecules may be used as hybridization probes, and may be labeled or unlabeled. In a preferred embodiment, the bound nucleic acid molecules are unlabeled.

In one embodiment, the nucleic acid molecule of the present invention is bound to a porous substrate, *e.g.*, a membrane, typically comprising nitrocellulose, nylon, or positively charged derivatized nylon. The nucleic acid molecule of the present invention can be used to detect a hybridizing nucleic acid molecule that is present within a labeled nucleic acid sample, *e.g.*, a sample of transcript-derived nucleic acids. In another embodiment, the nucleic acid molecule is bound to a solid substrate, including, without limitation, glass, amorphous silicon, crystalline silicon or plastics. Examples of plastics include, without limitation, polymethylacrylic, polyethylene, polypropylene, polyacrylate, polymethylmethacrylate, polyvinylchloride, polytetrafluoroethylene, polystyrene, polycarbonate, polyacetal, polysulfone, celluloseacetate, cellulosenitrate, nitrocellulose, or mixtures thereof. The solid substrate may be any shape, including rectangular, disk-like and spherical. In a preferred embodiment, the solid substrate is a microscope slide or slide-shaped substrate.

The nucleic acid molecule of the present invention can be attached covalently to a surface of the support substrate or applied to a derivatized surface in a chaotropic agent that facilitates denaturation and adherence by presumed noncovalent interactions, or some combination thereof. The nucleic acid molecule of the present invention can be bound to a substrate to which a plurality of other nucleic acids are concurrently bound, hybridization to each of the plurality of bound nucleic acids being separately detectable. At low density, *e.g.* on a porous membrane, these substrate-bound collections are typically denominated macroarrays; at higher density, typically on a solid support, such as glass, these substrate bound collections of plural nucleic acids are colloquially termed microarrays. As used herein, the term microarray includes arrays of all densities. It is, therefore, another aspect

of the invention to provide microarrays that comprise one or more of the nucleic acid molecules of the present invention.

In yet another embodiment, the invention is directed to single exon probes based on the CSNAs disclosed herein.

5 *Expression Vectors, Host Cells and Recombinant Methods of Producing Polypeptides*

Another aspect of the present invention provides vectors that comprise one or more of the isolated nucleic acid molecules of the present invention, and host cells in which such vectors have been introduced.

10 The vectors can be used, *inter alia*, for propagating the nucleic acid molecules of the present invention in host cells (cloning vectors), for shuttling the nucleic acid molecules of the present invention between host cells derived from disparate organisms (shuttle vectors), for inserting the nucleic acid molecules of the present invention into host cell chromosomes (insertion vectors), for expressing sense or antisense RNA transcripts of
15 the nucleic acid molecules of the present invention *in vitro* or within a host cell, and for expressing polypeptides encoded by the nucleic acid molecules of the present invention, alone or as fusion proteins with heterologous polypeptides (expression vectors). Vectors are by now well known in the art, and are described, *inter alia*, in Jones *et al.* (eds.), Vectors: Cloning Applications: Essential Techniques (Essential Techniques Series), John
20 Wiley & Son Ltd. (1998); Jones *et al.* (eds.), Vectors: Expression Systems: Essential Techniques (Essential Techniques Series), John Wiley & Son Ltd. (1998); Gacesa *et al.*, Vectors: Essential Data, John Wiley & Sons Ltd. (1995); Cid-Arregui (eds.), Viral Vectors: Basic Science and Gene Therapy, Eaton Publishing Co. (2000); Sambrook (2001), *supra*; Ausubel (1999), *supra*. Furthermore, a variety of vectors are available
25 commercially. Use of existing vectors and modifications thereof are well within the skill in the art. Thus, only basic features need be described here.

Nucleic acid sequences may be expressed by operatively linking them to an expression control sequence in an appropriate expression vector and employing that expression vector to transform an appropriate unicellular host. Expression control
30 sequences are sequences that control the transcription, post-transcriptional events and translation of nucleic acid sequences. Such operative linking of a nucleic acid sequence of this invention to an expression control sequence, of course, includes, if not already part of

the nucleic acid sequence, the provision of a translation initiation codon, ATG or GTG, in the correct reading frame upstream of the nucleic acid sequence.

A wide variety of host/expression vector combinations may be employed in expressing the nucleic acid sequences of this invention. Useful expression vectors, for example, may consist of segments of chromosomal, non-chromosomal and synthetic nucleic acid sequences.

In one embodiment, prokaryotic cells may be used with an appropriate vector. Prokaryotic host cells are often used for cloning and expression. In a preferred embodiment, prokaryotic host cells include *E. coli*, *Pseudomonas*, *Bacillus* and *Streptomyces*. In a preferred embodiment, bacterial host cells are used to express the nucleic acid molecules of the instant invention. Useful expression vectors for bacterial hosts include bacterial plasmids, such as those from *E. coli*, *Bacillus* or *Streptomyces*, including pBluescript, pGEX-2T, pUC vectors, col E1, pCR1, pBR322, pMB9 and their derivatives, wider host range plasmids, such as RP4, phage DNAs, *e.g.*, the numerous derivatives of phage lambda, *e.g.*, NM989, λ GT10 and λ GT11, and other phages, *e.g.*, M13 and filamentous single-stranded phage DNA. Where *E. coli* is used as host, selectable markers are, analogously, chosen for selectivity in gram negative bacteria: *e.g.*, typical markers confer resistance to antibiotics, such as ampicillin, tetracycline, chloramphenicol, kanamycin, streptomycin and zeocin; auxotrophic markers can also be used.

In other embodiments, eukaryotic host cells, such as yeast, insect, mammalian or plant cells, may be used. Yeast cells, typically *S. cerevisiae*, are useful for eukaryotic genetic studies, due to the ease of targeting genetic changes by homologous recombination and the ability to easily complement genetic defects using recombinantly expressed proteins. Yeast cells are useful for identifying interacting protein components, *e.g.* through use of a two-hybrid system. In a preferred embodiment, yeast cells are useful for protein expression. Vectors of the present invention for use in yeast will typically, but not invariably, contain an origin of replication suitable for use in yeast and a selectable marker that is functional in yeast. Yeast vectors include Yeast Integrating plasmids (*e.g.*, YIp5) and Yeast Replicating plasmids (the YRp and YEplac series plasmids), Yeast Centromere plasmids (the YCp series plasmids), Yeast Artificial Chromosomes (YACs) which are based on yeast linear plasmids, denoted YLp, pGPD-2, 2 μ plasmids and derivatives thereof, and improved shuttle vectors such as those described in Gietz *et al.*, *Gene*, 74:

527-34 (1988) (YIplac, YEplac and YCplac). Selectable markers in yeast vectors include a variety of auxotrophic markers, the most common of which are (in *Saccharomyces cerevisiae*) URA3, HIS3, LEU2, TRP1 and LYS2, which complement specific auxotrophic mutations, such as *ura3-52*, *his3-D1*, *leu2-D1*, *trp1-D1* and *lys2-201*.

- 5 Insect cells may be chosen for high efficiency protein expression. Where the host cells are from *Spodoptera frugiperda*, e.g., Sf9 and Sf21 cell lines, and expresSF™ cells (Protein Sciences Corp., Meriden, CT, USA), the vector replicative strategy is typically based upon the baculovirus life cycle. Typically, baculovirus transfer vectors are used to replace the wild-type AcMNPV polyhedrin gene with a heterologous gene of interest.
- 10 Sequences that flank the polyhedrin gene in the wild-type genome are positioned 5' and 3' of the expression cassette on the transfer vectors. Following co-transfection with AcMNPV DNA, a homologous recombination event occurs between these sequences resulting in a recombinant virus carrying the gene of interest and the polyhedrin or p10 promoter. Selection can be based upon visual screening for lacZ fusion activity.
- 15 The host cells may also be mammalian cells, which are particularly useful for expression of proteins intended as pharmaceutical agents, and for screening of potential agonists and antagonists of a protein or a physiological pathway. Mammalian vectors intended for autonomous extrachromosomal replication will typically include a viral origin, such as the SV40 origin (for replication in cell lines expressing the large T-antigen,
- 20 such as COS1 and COS7 cells), the papillomavirus origin, or the EBV origin for long term episomal replication (for use, e.g., in 293-EBNA cells, which constitutively express the EBV EBNA-1 gene product and adenovirus E1A). Vectors intended for integration, and thus replication as part of the mammalian chromosome, can, but need not, include an origin of replication functional in mammalian cells, such as the SV40 origin. Vectors
- 25 based upon viruses, such as adenovirus, adeno-associated virus, vaccinia virus, and various mammalian retroviruses, will typically replicate according to the viral replicative strategy. Selectable markers for use in mammalian cells include, but are not limited to, resistance to neomycin (G418), blasticidin, hygromycin and zeocin, and selection based upon the purine salvage pathway using HAT medium.
- 30 Expression in mammalian cells can be achieved using a variety of plasmids, including pSV2, pBC12BI, and p91023, as well as lytic virus vectors (e.g., vaccinia virus, adeno virus, and baculovirus), episomal virus vectors (e.g., bovine papillomavirus), and

retroviral vectors (*e.g.*, murine retroviruses). Useful vectors for insect cells include baculoviral vectors and pVL 941.

Plant cells can also be used for expression, with the vector replicon typically derived from a plant virus (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, 5 TMV) and selectable markers chosen for suitability in plants.

It is known that codon usage of different host cells may be different. For example, a plant cell and a human cell may exhibit a difference in codon preference for encoding a particular amino acid. As a result, human mRNA may not be efficiently translated in a plant, bacteria or insect host cell. Therefore, another embodiment of this invention is 10 directed to codon optimization. The codons of the nucleic acid molecules of the invention may be modified to resemble, as much as possible, genes naturally contained within the host cell without altering the amino acid sequence encoded by the nucleic acid molecule.

Any of a wide variety of expression control sequences may be used in these vectors to express the nucleic acid molecules of this invention. Such useful expression 15 control sequences include the expression control sequences associated with structural genes of the foregoing expression vectors. Expression control sequences that control transcription include, *e.g.*, promoters, enhancers and transcription termination sites. Expression control sequences in eukaryotic cells that control post-transcriptional events include splice donor and acceptor sites and sequences that modify the half-life of the 20 transcribed RNA, *e.g.*, sequences that direct poly(A) addition or binding sites for RNA-binding proteins. Expression control sequences that control translation include ribosome binding sites, sequences which direct targeted expression of the polypeptide to or within particular cellular compartments, and sequences in the 5' and 3' untranslated regions that modify the rate or efficiency of translation.

25 Examples of useful expression control sequences for a prokaryote, *e.g.*, *E. coli*, will include a promoter, often a phage promoter, such as phage lambda pL promoter, the trc promoter, a hybrid derived from the trp and lac promoters, the bacteriophage T7 promoter (in *E. coli* cells engineered to express the T7 polymerase), the TAC or TRC system, the major operator and promoter regions of phage lambda, the control regions of 30 fd coat protein, and the araBAD operon. Prokaryotic expression vectors may further include transcription terminators, such as the aspA terminator, and elements that facilitate translation, such as a consensus ribosome binding site and translation termination codon, Schomer *et al.*, *Proc. Natl. Acad. Sci. USA* 83: 8506-8510 (1986).

Expression control sequences for yeast cells, typically *S. cerevisiae*, will include a yeast promoter, such as the CYC1 promoter, the GAL1 promoter, the GAL10 promoter, ADH1 promoter, the promoters of the yeast α -mating system, or the GPD promoter, and will typically have elements that facilitate transcription termination, such as the
5 transcription termination signals from the CYC1 or ADH1 gene.

Expression vectors useful for expressing proteins in mammalian cells will include a promoter active in mammalian cells. These promoters include, but are not limited to, those derived from mammalian viruses, such as the enhancer-promoter sequences from the immediate early gene of the human cytomegalovirus (CMV), the enhancer-promoter
10 sequences from the Rous sarcoma virus long terminal repeat (RSV LTR), the enhancer-promoter from SV40 and the early and late promoters of adenovirus. Other expression control sequences include the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase. Other expression control sequences include those from the gene comprising the CSNA of interest. Often, expression is enhanced by
15 incorporation of polyadenylation sites, such as the late SV40 polyadenylation site and the polyadenylation signal and transcription termination sequences from the bovine growth hormone (BGH) gene, and ribosome binding sites. Furthermore, vectors can include introns, such as intron II of rabbit β -globin gene and the SV40 splice elements.

Preferred nucleic acid vectors also include a selectable or amplifiable marker gene
20 and means for amplifying the copy number of the gene of interest. Such marker genes are well known in the art. Nucleic acid vectors may also comprise stabilizing sequences (*e.g.*, ori- or ARS-like sequences and telomere-like sequences), or may alternatively be designed to favor directed or non-directed integration into the host cell genome. In a preferred embodiment, nucleic acid sequences of this invention are inserted in frame into an
25 expression vector that allows a high level expression of an RNA which encodes a protein comprising the encoded nucleic acid sequence of interest. Nucleic acid cloning and sequencing methods are well known to those of skill in the art and are described in an assortment of laboratory manuals, including Sambrook (1989), *supra*, Sambrook (2000), *supra*; Ausubel (1992), *supra*; and Ausubel (1999), *supra*. Product information from
30 manufacturers of biological, chemical and immunological reagents also provide useful information.

Expression vectors may be either constitutive or inducible. Inducible vectors include either naturally inducible promoters, such as the *trc* promoter, which is regulated

by the lac operon, and the pL promoter, which is regulated by tryptophan, the MMTV-LTR promoter, which is inducible by dexamethasone, or can contain synthetic promoters and/or additional elements that confer inducible control on adjacent promoters. Examples of inducible synthetic promoters are the hybrid Plac/ara-1 promoter and the

5 PLtetO-1 promoter. The PLtetO-1 promoter takes advantage of the high expression levels from the PL promoter of phage lambda, but replaces the lambda repressor sites with two copies of operator 2 of the Tn10 tetracycline resistance operon, causing this promoter to be tightly repressed by the Tet repressor protein and induced in response to tetracycline (Tc) and Tc derivatives such as anhydrotetracycline. Vectors may also be inducible

10 because they contain hormone response elements, such as the glucocorticoid response element (GRE) and the estrogen response element (ERE), which can confer hormone inducibility where vectors are used for expression in cells having the respective hormone receptors. To reduce background levels of expression, elements responsive to ecdysone, an insect hormone, can be used instead, with coexpression of the ecdysone receptor.

15 In one embodiment of the invention, expression vectors can be designed to fuse the expressed polypeptide to small protein tags that facilitate purification and/or visualization. Such tags include a polyhistidine tag that facilitates purification of the fusion protein by immobilized metal affinity chromatography, for example using NiNTA resin (Qiagen Inc., Valencia, CA, USA) or TALON™ resin (cobalt immobilized affinity chromatography

20 medium, Clontech Labs, Palo Alto, CA, USA). The fusion protein can include a chitin-binding tag and self-excising intein, permitting chitin-based purification with self-removal of the fused tag (IMPACT™ system, New England Biolabs, Inc., Beverly, MA, USA). Alternatively, the fusion protein can include a calmodulin-binding peptide tag, permitting purification by calmodulin affinity resin (Stratagene, La Jolla, CA, USA), or a specifically

25 excisable fragment of the biotin carboxylase carrier protein, permitting purification of *in vivo* biotinylated protein using an avidin resin and subsequent tag removal (Promega, Madison, WI, USA). As another useful alternative, the polypeptides of the present invention can be expressed as a fusion to glutathione-S-transferase, the affinity and specificity of binding to glutathione permitting purification using glutathione affinity

30 resins, such as Glutathione-Superflow Resin (Clontech Laboratories, Palo Alto, CA, USA), with subsequent elution with free glutathione. Other tags include, for example, the Xpress epitope, detectable by anti-Xpress antibody (Invitrogen, Carlsbad, CA, USA), a myc tag, detectable by anti-myc tag antibody, the V5 epitope, detectable by anti-V5

antibody (Invitrogen, Carlsbad, CA, USA), FLAG® epitope, detectable by anti-FLAG® antibody (Stratagene, La Jolla, CA, USA), and the HA epitope, detectable by anti-HA antibody.

For secretion of expressed polypeptides, vectors can include appropriate sequences
5 that encode secretion signals, such as leader peptides. For example, the pSecTag2 vectors (Invitrogen, Carlsbad, CA, USA) are 5.2 kb mammalian expression vectors that carry the secretion signal from the V-J2-C region of the mouse Ig kappa-chain for efficient secretion of recombinant proteins from a variety of mammalian cell lines.

Expression vectors can also be designed to fuse proteins encoded by the
10 heterologous nucleic acid insert to polypeptides that are larger than purification and/or identification tags. Useful protein fusions include those that permit display of the encoded protein on the surface of a phage or cell, fusions to intrinsically fluorescent proteins, such as those that have a green fluorescent protein (GFP)-like chromophore, fusions to the IgG Fc region, and fusions for use in two hybrid systems.

15 Vectors for phage display fuse the encoded polypeptide to, e.g., the gene III protein (pIII) or gene VIII protein (pVIII) for display on the surface of filamentous phage, such as M13. *See* Barbas *et al.*, Phage Display: A Laboratory Manual, Cold Spring Harbor Laboratory Press (2001); Kay *et al.* (eds.), Phage Display of Peptides and Proteins: A Laboratory Manual, Academic Press, Inc., (1996); Abelson *et al.* (eds.), Combinatorial Chemistry (Methods in Enzymology, Vol. 267) Academic Press (1996). Vectors for yeast
20 display, e.g. the pYD1 yeast display vector (Invitrogen, Carlsbad, CA, USA), use the α -agglutinin yeast adhesion receptor to display recombinant protein on the surface of *S. cerevisiae*. Vectors for mammalian display, e.g., the pDisplay™ vector (Invitrogen, Carlsbad, CA, USA), target recombinant proteins using an N-terminal cell surface
25 targeting signal and a C-terminal transmembrane anchoring domain of platelet derived growth factor receptor.

A wide variety of vectors now exist that fuse proteins encoded by heterologous nucleic acids to the chromophore of the substrate-independent, intrinsically fluorescent green fluorescent protein from *Aequorea victoria* ("GFP") and its variants. The GFP-like
30 chromophore can be selected from GFP-like chromophores found in naturally occurring proteins, such as *A. victoria* GFP (GenBank accession number AAA27721), *Renilla reniformis* GFP, FP583 (GenBank accession no. AF168419) (DsRed), FP593 (AF272711), FP483 (AF168420), FP484 (AF168424), FP595 (AF246709), FP486 (AF168421), FP538

(AF168423), and FP506 (AF168422), and need include only so much of the native protein as is needed to retain the chromophore's intrinsic fluorescence. Methods for determining the minimal domain required for fluorescence are known in the art. *See Li et al., J. Biol. Chem.* 272: 28545-28549 (1997). Alternatively, the GFP-like chromophore can be
5 selected from GFP-like chromophores modified from those found in nature. The methods for engineering such modified GFP-like chromophores and testing them for fluorescence activity, both alone and as part of protein fusions, are well known in the art. *See Heim et al., Curr. Biol.* 6: 178-182 (1996) and Palm *et al., Methods Enzymol.* 302: 378-394 (1999). A variety of such modified chromophores are now commercially available and can readily
10 be used in the fusion proteins of the present invention. These include EGFP ("enhanced GFP"), EBFP ("enhanced blue fluorescent protein"), BFP2, EYFP ("enhanced yellow fluorescent protein"), ECFP ("enhanced cyan fluorescent protein") or Citrine. EGFP (*see, e.g., Cormack et al., Gene* 173: 33-38 (1996); U.S. Patent Nos. 6,090,919 and 5,804,387, the disclosures of which are incorporated herein by reference in their entireties) is found
15 on a variety of vectors, both plasmid and viral, which are available commercially (Clontech Labs, Palo Alto, CA, USA); EBFP is optimized for expression in mammalian cells whereas BFP2, which retains the original jellyfish codons, can be expressed in bacteria (*see, e.g., Heim et al., Curr. Biol.* 6: 178-182 (1996) and Cormack *et al., Gene* 173: 33-38 (1996)). Vectors containing these blue-shifted variants are available from
20 Clontech Labs (Palo Alto, CA, USA). Vectors containing EYFP, ECFP (*see, e.g., Heim et al., Curr. Biol.* 6: 178-182 (1996); Miyawaki *et al., Nature* 388: 882-887 (1997)) and Citrine (*see, e.g., Heikal et al., Proc. Natl. Acad. Sci. USA* 97: 11996-12001 (2000)) are also available from Clontech Labs. The GFP-like chromophore can also be drawn from other modified GFPs, including those described in U.S. Patent Nos. 6,124,128; 6,096,865;
25 6,090,919; 6,066,476; 6,054,321; 6,027,881; 5,968,750; 5,874,304; 5,804,387; 5,777,079; 5,741,668; and 5,625,048, the disclosures of which are incorporated herein by reference in their entireties. *See also Conn (ed.), Green Fluorescent Protein* (Methods in Enzymology, Vol. 302), Academic Press, Inc. (1999); Yang, *et al., J Biol Chem*, 273: 8212-6 (1998); Bevis *et al., Nature Biotechnology*, 20:83-7 (2002). The GFP-like chromophore of each
30 of these GFP variants can usefully be included in the fusion proteins of the present invention.

Fusions to the IgG Fc region increase serum half-life of protein pharmaceutical products through interaction with the FcRn receptor (also denominated the FcRp receptor

and the Brambell receptor, FcRb), further described in International Patent Application Nos. WO 97/43316, WO 97/34631, WO 96/32478, and WO 96/18412, the disclosures of which are incorporated herein by reference in their entireties.

For long-term, high-yield recombinant production of the polypeptides of the present invention, stable expression is preferred. Stable expression is readily achieved by integration into the host cell genome of vectors having selectable markers, followed by selection of these integrants. Vectors such as pUB6/V5-His A, B, and C (Invitrogen, Carlsbad, CA, USA) are designed for high-level stable expression of heterologous proteins in a wide range of mammalian tissue types and cell lines. pUB6/V5-His uses the promoter/enhancer sequence from the human ubiquitin C gene to drive expression of recombinant proteins: expression levels in 293, CHO, and NIH3T3 cells are comparable to levels from the CMV and human EF-1a promoters. The bsd gene permits rapid selection of stably transfected mammalian cells with the potent antibiotic blasticidin.

Replication incompetent retroviral vectors, typically derived from Moloney murine leukemia virus, also are useful for creating stable transfectants having integrated provirus. The highly efficient transduction machinery of retroviruses, coupled with the availability of a variety of packaging cell lines such as RetroPack™ PT 67, EcoPack2™-293, AmphoPack-293, and GP2-293 cell lines (all available from Clontech Laboratories, Palo Alto, CA, USA) allow a wide host range to be infected with high efficiency; varying the multiplicity of infection readily adjusts the copy number of the integrated provirus.

Of course, not all vectors and expression control sequences will function equally well to express the nucleic acid molecules of this invention. Neither will all hosts function equally well with the same expression system. However, one of skill in the art may make a selection among these vectors, expression control sequences and hosts without undue experimentation and without departing from the scope of this invention. For example, in selecting a vector, the host must be considered because the vector must be replicated in it. The vector's copy number, the ability to control that copy number, the ability to control integration, if any, and the expression of any other proteins encoded by the vector, such as an antibiotic or other selection marker, should also be considered. The present invention further includes host cells comprising the vectors of the present invention, either present episomally within the cell or integrated, in whole or in part, into the host cell chromosome. Among other considerations, some of which are described above, a host cell strain may be chosen for its ability to process the expressed polypeptide in the desired fashion. Such

post-translational modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation, and it is an aspect of the present invention to provide CSPs with such post-translational modifications.

5 In selecting an expression control sequence, a variety of factors should also be considered. These include, for example, the relative strength of the sequence, its controllability, and its compatibility with the nucleic acid molecules of this invention, particularly with regard to potential secondary structures. Unicellular hosts should be selected by consideration of their compatibility with the chosen vector, the toxicity of the
10 product coded for by the nucleic acid sequences of this invention, their secretion characteristics, their ability to fold the polypeptide correctly, their fermentation or culture requirements, and the ease of purification from them of the products coded for by the nucleic acid molecules of this invention.

 The recombinant nucleic acid molecules and more particularly, the expression
15 vectors of this invention may be used to express the polypeptides of this invention as recombinant polypeptides in a heterologous host cell. The polypeptides of this invention may be full-length or less than full-length polypeptide fragments recombinantly expressed from the nucleic acid molecules according to this invention. Such polypeptides include analogs, derivatives and muteins that may or may not have biological activity.

20 Vectors of the present invention will also often include elements that permit *in vitro* transcription of RNA from the inserted heterologous nucleic acid. Such vectors typically include a phage promoter, such as that from T7, T3, or SP6, flanking the nucleic acid insert. Often two different such promoters flank the inserted nucleic acid, permitting separate *in vitro* production of both sense and antisense strands.

25 Transformation and other methods of introducing nucleic acids into a host cell (*e.g.*, conjugation, protoplast transformation or fusion, transfection, electroporation, liposome delivery, membrane fusion techniques, high velocity DNA-coated pellets, viral infection and protoplast fusion) can be accomplished by a variety of methods which are well known in the art (*See*, for instance, Ausubel, *supra*, and Sambrook *et al.*, *supra*).

30 Bacterial, yeast, plant or mammalian cells are transformed or transfected with an expression vector, such as a plasmid, a cosmid, or the like, wherein the expression vector comprises the nucleic acid of interest. Alternatively, the cells may be infected by a viral expression vector comprising the nucleic acid of interest. Depending upon the host cell,

vector, and method of transformation used, transient or stable expression of the polypeptide will be constitutive or inducible. One having ordinary skill in the art will be able to decide whether to express a polypeptide transiently or stably, and whether to express the protein constitutively or inducibly.

5 A wide variety of unicellular host cells are useful in expressing the DNA sequences of this invention. These hosts may include well known eukaryotic and prokaryotic hosts, such as strains of, fungi, yeast, insect cells such as *Spodoptera frugiperda* (SF9), animal cells such as CHO, as well as plant cells in tissue culture. Representative examples of appropriate host cells include, but are not limited to, bacterial
10 cells, such as *E. coli*, *Caulobacter crescentus*, *Streptomyces* species, and *Salmonella typhimurium*; yeast cells, such as *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Pichia pastoris*, *Pichia methanolica*; insect cell lines, such as those from *Spodoptera frugiperda*, e.g., Sf9 and Sf21 cell lines, and expresSFTM cells (Protein Sciences Corp., Meriden, CT, USA), *Drosophila* S2 cells, and *Trichoplusia ni* High Five® Cells
15 (Invitrogen, Carlsbad, CA, USA); and mammalian cells. Typical mammalian cells include BHK cells, BSC 1 cells, BSC 40 cells, BMT 10 cells, VERO cells, COS1 cells, COS7 cells, Chinese hamster ovary (CHO) cells, 3T3 cells, NIH 3T3 cells, 293 cells, HEPG2 cells, HeLa cells, L cells, MDCK cells, HEK293 cells, WI38 cells, murine ES cell lines (e.g., from strains 129/SV, C57/BL6, DBA-1, 129/SVJ), K562 cells, Jurkat cells, and
20 BW5147 cells. Other mammalian cell lines are well known and readily available from the American Type Culture Collection (ATCC) (Manassas, VA, USA) and the National Institute of General Medical Sciences (NIGMS) Human Genetic Cell Repository at the Coriell Cell Repositories (Camden, NJ, USA). Cells or cell lines derived from colon are particularly preferred because they may provide a more native post-translational
25 processing. Particularly preferred are human colon cells.

Particular details of the transfection, expression and purification of recombinant proteins are well documented and are understood by those of skill in the art. Further details on the various technical aspects of each of the steps used in recombinant production of foreign genes in bacterial cell expression systems can be found in a number
30 of texts and laboratory manuals in the art. See, e.g., Ausubel (1992), *supra*, Ausubel (1999), *supra*, Sambrook (1989), *supra*, and Sambrook (2001), *supra*.

Methods for introducing the vectors and nucleic acid molecules of the present invention into the host cells are well known in the art; the choice of technique will depend primarily upon the specific vector to be introduced and the host cell chosen.

Nucleic acid molecules and vectors may be introduced into prokaryotes, such as *E. coli*, in a number of ways. For instance, phage lambda vectors will typically be packaged using a packaging extract (e.g., Gigapack® packaging extract, Stratagene, La Jolla, CA, USA), and the packaged virus used to infect *E. coli*.

Plasmid vectors will typically be introduced into chemically competent or electrocompetent bacterial cells. *E. coli* cells can be rendered chemically competent by treatment, e.g., with CaCl_2 , or a solution of Mg^{2+} , Mn^{2+} , Ca^{2+} , Rb^+ or K^+ , dimethyl sulfoxide, dithiothreitol, and hexamine cobalt (III), Hanahan, *J. Mol. Biol.* 166(4):557-80 (1983), and vectors introduced by heat shock. A wide variety of chemically competent strains are also available commercially (e.g., Epicurian Coli® XL10-Gold® Ultracompetent Cells (Stratagene, La Jolla, CA, USA); DH5α competent cells (Clontech Laboratories, Palo Alto, CA, USA); and TOP10 Chemically Competent *E. coli* Kit (Invitrogen, Carlsbad, CA, USA)). Bacterial cells can be rendered electrocompetent to take up exogenous DNA by electroporation by various pre-pulse treatments; vectors are introduced by electroporation followed by subsequent outgrowth in selected media. An extensive series of protocols is provided by BioRad (Richmond, CA, USA).

Vectors can be introduced into yeast cells by spheroplasting, treatment with lithium salts, electroporation, or protoplast fusion. Spheroplasts are prepared by the action of hydrolytic enzymes such as a snail-gut extract, usually denoted Glusulase or Zymolyase, or an enzyme from *Arthrobacter luteus* to remove portions of the cell wall in the presence of osmotic stabilizers, typically 1 M sorbitol. DNA is added to the spheroplasts, and the mixture is co-precipitated with a solution of polyethylene glycol (PEG) and Ca^{2+} . Subsequently, the cells are resuspended in a solution of sorbitol, mixed with molten agar and then layered on the surface of a selective plate containing sorbitol.

For lithium-mediated transformation, yeast cells are treated with lithium acetate to permeabilize the cell wall, DNA is added and the cells are co-precipitated with PEG. The cells are exposed to a brief heat shock, washed free of PEG and lithium acetate, and subsequently spread on plates containing ordinary selective medium. Increased frequencies of transformation are obtained by using specially-prepared single-stranded

carrier DNA and certain organic solvents. Schiestl *et al.*, *Curr. Genet.* 16(5-6): 339-46 (1989).

For electroporation, freshly-grown yeast cultures are typically washed, suspended in an osmotic protectant, such as sorbitol, mixed with DNA, and the cell suspension pulsed in an electroporation device. Subsequently, the cells are spread on the surface of plates containing selective media. Becker *et al.*, *Methods Enzymol.* 194: 182-187 (1991). The efficiency of transformation by electroporation can be increased over 100-fold by using PEG, single-stranded carrier DNA and cells that are in late log-phase of growth. Larger constructs, such as YACs, can be introduced by protoplast fusion.

Mammalian and insect cells can be directly infected by packaged viral vectors, or transfected by chemical or electrical means. For chemical transfection, DNA can be coprecipitated with CaPO_4 or introduced using liposomal and nonliposomal lipid-based agents. Commercial kits are available for CaPO_4 transfection (CalPhos™ Mammalian Transfection Kit, Clontech Laboratories, Palo Alto, CA, USA), and lipid-mediated transfection can be practiced using commercial reagents, such as LIPOFECTAMINE™ 2000, LIPOFECTAMINE™ Reagent, CELLFECTIN® Reagent, and LIPOFECTIN® Reagent (Invitrogen, Carlsbad, CA, USA), DOTAP Liposomal Transfection Reagent, FuGENE 6, X-tremeGENE Q2, DOSPER, (Roche Molecular Biochemicals, Indianapolis, IN USA), Effectene™, PolyFect®, Superfect® (Qiagen, Inc., Valencia, CA, USA). Protocols for electroporating mammalian cells can be found in, for example, ; Norton *et al.* (eds.), Gene Transfer Methods: Introducing DNA into Living Cells and Organisms, BioTechniques Books, Eaton Publishing Co. (2000). Other transfection techniques include transfection by particle bombardment and microinjection. See, e.g., Cheng *et al.*, *Proc. Natl. Acad. Sci. USA* 90(10): 4455-9 (1993); Yang *et al.*, *Proc. Natl. Acad. Sci. USA* 87(24): 9568-72 (1990).

Production of the recombinantly produced proteins of the present invention can optionally be followed by purification.

Purification of recombinantly expressed proteins is now well within the skill in the art and thus need not be detailed here. See, e.g., Thorner *et al.* (eds.), Applications of Chimeric Genes and Hybrid Proteins, Part A: Gene Expression and Protein Purification (Methods in Enzymology, Vol. 326), Academic Press (2000); Harbin (ed.), Cloning, Gene Expression and Protein Purification : Experimental Procedures and Process Rationale, Oxford Univ. Press (2001); Marshak *et al.*, Strategies for Protein Purification and

Characterization: A Laboratory Course Manual, Cold Spring Harbor Laboratory Press (1996); and Roe (ed.), Protein Purification Applications, Oxford University Press (2001).

Briefly, however, if purification tags have been fused through use of an expression vector that appends such tags, purification can be effected, at least in part, by means
5 appropriate to the tag, such as use of immobilized metal affinity chromatography for polyhistidine tags. Other techniques common in the art include ammonium sulfate fractionation, immunoprecipitation, fast protein liquid chromatography (FPLC), high performance liquid chromatography (HPLC), and preparative gel electrophoresis.

10 Polypeptides, including Fragments Muteins, Homologous Proteins, Allelic Variants, Analogs and Derivatives

Another aspect of the invention relates to polypeptides encoded by the nucleic acid molecules described herein. In a preferred embodiment, the polypeptide is a colon specific polypeptide (CSP). In an even more preferred embodiment, the polypeptide comprises an amino acid sequence of SEQ ID NO:96-237 or is derived from a polypeptide
15 having the amino acid sequence of SEQ ID NO: 96-237. A polypeptide as defined herein may be produced recombinantly, as discussed *supra*, may be isolated from a cell that naturally expresses the protein, or may be chemically synthesized following the teachings of the specification and using methods well known to those having ordinary skill in the art.

Polypeptides of the present invention may also comprise a part or fragment of a
20 CSP. In a preferred embodiment, the fragment is derived from a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 96-237. Polypeptides of the present invention comprising a part or fragment of an entire CSP may or may not be CSPs. For example, a full-length polypeptide may be colon-specific, while a fragment thereof may be found in other tissues as well as in colon. A polypeptide that is
25 not a CSP, whether it is a fragment, analog, mutein, homologous protein or derivative, is nevertheless useful, especially for immunizing animals to prepare anti-CSP antibodies. In a preferred embodiment, the part or fragment is a CSP. Methods of determining whether a polypeptide of the present invention is a CSP are described *infra*.

Polypeptides of the present invention comprising fragments of at least 6
30 contiguous amino acids are also useful in mapping B cell and T cell epitopes of the reference protein. See, e.g., Geysen *et al.*, *Proc. Natl. Acad. Sci. USA* 81: 3998-4002 (1984) and U.S. Patent Nos. 4,708,871 and 5,595,915, the disclosures of which are

incorporated herein by reference in their entireties. Because the fragment need not itself be immunogenic, part of an immunodominant epitope, nor even recognized by native antibody, to be useful in such epitope mapping, all fragments of at least 6 amino acids of a polypeptide of the present invention have utility in such a study.

5 Polypeptides of the present invention comprising fragments of at least 8 contiguous amino acids, often at least 15 contiguous amino acids, are useful as immunogens for raising antibodies that recognize polypeptides of the present invention. See, e.g., Lerner, *Nature* 299: 592-596 (1982); Shinnick *et al.*, *Annu. Rev. Microbiol.* 37: 425-46 (1983); Sutcliffe *et al.*, *Science* 219: 660-6 (1983). As further described in the
10 above-cited references, virtually all 8-mers, conjugated to a carrier, such as a protein, prove immunogenic and are capable of eliciting antibody for the conjugated peptide; accordingly, all fragments of at least 8 amino acids of the polypeptides of the present invention have utility as immunogens.

Polypeptides comprising fragments of at least 8, 9, 10 or 12 contiguous amino
15 acids are also useful as competitive inhibitors of binding of the entire polypeptide, or a portion thereof, to antibodies (as in epitope mapping), and to natural binding partners, such as subunits in a multimeric complex or to receptors or ligands of the subject protein; this competitive inhibition permits identification and separation of molecules that bind specifically to the polypeptide of interest. See U.S. Patent Nos. 5,539,084 and 5,783,674,
20 incorporated herein by reference in their entireties.

The polypeptide of the present invention thus preferably is at least 6 amino acids in length, typically at least 8, 9, 10 or 12 amino acids in length, and often at least 15 amino acids in length. Often, the polypeptide of the present invention is at least 20 amino acids in length, even 25 amino acids, 30 amino acids, 35 amino acids, or 50 amino acids or more
25 in length. Of course, larger polypeptides having at least 75 amino acids, 100 amino acids, or even 150 amino acids are also useful, and at times preferred.

One having ordinary skill in the art can produce fragments by truncating the nucleic acid molecule, e.g., a CSNA, encoding the polypeptide and then expressing it recombinantly. Alternatively, one can produce a fragment by chemically synthesizing a
30 portion of the full-length polypeptide. One may also produce a fragment by enzymatically cleaving either a recombinant polypeptide or an isolated naturally occurring polypeptide. Methods of producing polypeptide fragments are well known in the art. See, e.g., Sambrook (1989), *supra*; Sambrook (2001), *supra*; Ausubel (1992), *supra*; and Ausubel

(1999), *supra*. In one embodiment, a polypeptide comprising only a fragment, preferably a fragment of a CSP, may be produced by chemical or enzymatic cleavage of a CSP polypeptide. In a preferred embodiment, a polypeptide fragment is produced by expressing a nucleic acid molecule of the present invention encoding a fragment, preferably of a CSP, in a host cell.

Polypeptides of the present invention are also inclusive of mutants, fusion proteins, homologous proteins and allelic variants.

A mutant protein, or mutein, may have the same or different properties compared to a naturally occurring polypeptide and comprises at least one amino acid insertion, duplication, deletion, rearrangement or substitution compared to the amino acid sequence of a native polypeptide. Small deletions and insertions can often be found that do not alter the function of a protein. Muteins may or may not be colon-specific. Preferably, the mutein is colon-specific. More preferably the mutein is a polypeptide that comprises at least one amino acid insertion, duplication, deletion, rearrangement or substitution compared to the amino acid sequence of SEQ ID NO: 96-237. Accordingly, in a preferred embodiment, the mutein is one that exhibits at least 50% sequence identity, more preferably at least 60% sequence identity, even more preferably at least 70%, yet more preferably at least 80% sequence identity to a CSP comprising an amino acid sequence of SEQ ID NO: 96-237. In a yet more preferred embodiment, the mutein exhibits at least 85%, more preferably 90%, even more preferably 95% or 96%, and yet more preferably at least 97%, 98%, 99% or 99.5% sequence identity to a CSP comprising an amino acid sequence of SEQ ID NO: 96-237.

A mutein may be produced by isolation from a naturally occurring mutant cell, tissue or organism. A mutein may be produced by isolation from a cell, tissue or organism that has been experimentally mutagenized. Alternatively, a mutein may be produced by chemical manipulation of a polypeptide, such as by altering the amino acid residue to another amino acid residue using synthetic or semi-synthetic chemical techniques. In a preferred embodiment, a mutein is produced from a host cell comprising a mutated nucleic acid molecule compared to the naturally occurring nucleic acid molecule. For instance, one may produce a mutein of a polypeptide by introducing one or more mutations into a nucleic acid molecule of the invention and then expressing it recombinantly. These mutations may be targeted, in which particular encoded amino acids are altered, or may be untargeted, in which random encoded amino acids within the polypeptide are altered.

Mutagens with random amino acid alterations can be screened for a particular biological activity or property, particularly whether the polypeptide is colon-specific, as described below. Multiple random mutations can be introduced into the gene by methods well known to the art, *e.g.*, by error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, *in vivo* mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis and site-specific mutagenesis. Methods of producing mutagens with targeted or random amino acid alterations are well known in the art. *See, e.g.*, Sambrook (1989), *supra*; Sambrook (2001), *supra*; Ausubel (1992), *supra*; and Ausubel (1999), as well as U.S. Patent No. 5,223,408, which is herein incorporated by reference in its entirety.

The invention also contemplates polypeptides that are homologous to a polypeptide of the invention. In a preferred embodiment, the polypeptide is homologous to a CSP. In an even more preferred embodiment, the polypeptide is homologous to a CSP selected from the group having an amino acid sequence of SEQ ID NO: 96-237. By homologous polypeptide it is meant one that exhibits significant sequence identity to a CSP, preferably a CSP having an amino acid sequence of SEQ ID NO: 96-237. By significant sequence identity it is meant that the homologous polypeptide exhibits at least 50% sequence identity, more preferably at least 60% sequence identity, even more preferably at least 70%, yet more preferably at least 80% sequence identity to a CSP comprising an amino acid sequence of SEQ ID NO: 96-237. More preferred are homologous polypeptides exhibiting at least 85%, more preferably 90%, even more preferably 95% or 96%, and yet more preferably at least 97% or 98% sequence identity to a CSP comprising an amino acid sequence of SEQ ID NO: 96-237. Most preferably, the homologous polypeptide exhibits at least 99%, more preferably 99.5%, even more preferably 99.6%, 99.7%, 99.8% or 99.9% sequence identity to a CSP comprising an amino acid sequence of SEQ ID NO: 96-237. In a preferred embodiment, the amino acid substitutions of the homologous polypeptide are conservative amino acid substitutions as discussed *supra*.

Homologous polypeptides of the present invention also comprise polypeptide encoded by a nucleic acid molecule that selectively hybridizes to a CSNA or an antisense sequence thereof. In this embodiment, it is preferred that the homologous polypeptide be encoded by a nucleic acid molecule that hybridizes to a CSNA under low stringency, moderate stringency or high stringency conditions, as defined herein. More preferred is a

homologous polypeptide encoded by a nucleic acid sequence which hybridizes to a CSNA selected from the group consisting of SEQ ID NO: 1-95 or a homologous polypeptide encoded by a nucleic acid molecule that hybridizes to a nucleic acid molecule that encodes a CSP, preferably a CSP of SEQ ID NO:96-237 under low stringency, moderate
5 stringency or high stringency conditions, as defined herein.

Homologous polypeptides of the present invention may be naturally occurring and derived from another species, especially one derived from another primate, such as chimpanzee, gorilla, rhesus macaque, or baboon, wherein the homologous polypeptide comprises an amino acid sequence that exhibits significant sequence identity to that of
10 SEQ ID NO: 96-237. The homologous polypeptide may also be a naturally occurring polypeptide from a human, when the CSP is a member of a family of polypeptides. The homologous polypeptide may also be a naturally occurring polypeptide derived from a non-primate, mammalian species, including without limitation, domesticated species, *e.g.*, dog, cat, mouse, rat, rabbit, guinea pig, hamster, cow, horse, goat or pig. The homologous
15 polypeptide may also be a naturally occurring polypeptide derived from a non-mammalian species, such as birds or reptiles. The naturally occurring homologous protein may be isolated directly from humans or other species. Alternatively, the nucleic acid molecule encoding the naturally occurring homologous polypeptide may be isolated and used to express the homologous polypeptide recombinantly. The homologous polypeptide may
20 also be one that is experimentally produced by random mutation of a nucleic acid molecule and subsequent expression of the nucleic acid molecule. Alternatively, the homologous polypeptide may be one that is experimentally produced by directed mutation of one or more codons to alter the encoded amino acid of a CSP. In a preferred embodiment, the homologous polypeptide encodes a polypeptide that is a CSP.

25 Relatedness of proteins can also be characterized using a second functional test, such as the ability of a first protein competitively to inhibit the binding of a second protein to an antibody. It is, therefore, another aspect of the present invention to provide isolated polypeptides not only identical in sequence to those described with particularity herein, but also to provide isolated polypeptides ("cross-reactive proteins") that competitively
30 inhibit the binding of antibodies to all or to a portion of the isolated polypeptides of the present invention. Such competitive inhibition can readily be determined using immunoassays well known in the art.

As discussed above, single nucleotide polymorphisms (SNPs) occur frequently in eukaryotic genomes, and the sequence determined from one individual of a species may differ from other allelic forms present within the population. Thus, polypeptides of the present invention are also inclusive of those encoded by an allelic variant of a nucleic acid molecule encoding a CSP. In this embodiment, it is preferred that the polypeptide be encoded by an allelic variant of a gene that encodes a polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO: 96-237. More preferred is that the polypeptide be encoded by an allelic variant of a gene that has the nucleic acid sequence selected from the group consisting of SEQ ID NO: 1-95.

Polypeptides of the present invention are also inclusive of derivative polypeptides encoded by a nucleic acid molecule according to the instant invention. In this embodiment, it is preferred that the polypeptide be a CSP. Also preferred are derivative polypeptides having an amino acid sequence selected from the group consisting of SEQ ID NO: 96-237 and which has been acetylated, carboxylated, phosphorylated, glycosylated, ubiquitinated or post-translationally modified in another manner. In another preferred embodiment, the derivative has been labeled with, *e.g.*, radioactive isotopes such as ^{125}I , ^{32}P , ^{35}S , and ^3H . In another preferred embodiment, the derivative has been labeled with fluorophores, chemiluminescent agents, enzymes, and antigens that can serve as specific binding pair members for a labeled ligand.

Polypeptide modifications are well known to those of skill and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as, for instance Creighton, Protein Structure and Molecular Properties, 2nd ed., W. H. Freeman and Company (1993). Many detailed reviews are available on this subject, such as, for example, those provided by Wold, in Johnson (ed.), Posttranslational Covalent Modification of Proteins, pgs. 1-12, Academic Press (1983); Seifter *et al.*, *Meth. Enzymol.* 182: 626-646 (1990) and Rattan *et al.*, *Ann. N.Y. Acad. Sci.* 663: 48-62 (1992).

One may determine whether a polypeptide of the invention is likely to be post-translationally modified by analyzing the sequence of the polypeptide to determine if there are peptide motifs indicative of sites for post-translational modification. There are a number of computer programs that permit prediction of post-translational modifications.

See, e.g., expasy.org (accessed November 11, 2002) of the world wide web, which includes PSORT, for prediction of protein sorting signals and localization sites, SignalP, for prediction of signal peptide cleavage sites, MITOPROT and Predotar, for prediction of mitochondrial targeting sequences, NetOGlyc, for prediction of type O-glycosylation sites in mammalian proteins, big-PI Predictor and DGPI, for prediction of prenylation-anchor and cleavage sites, and NetPhos, for prediction of Ser, Thr and Tyr phosphorylation sites in eukaryotic proteins. Other computer programs, such as those included in GCG, also may be used to determine post-translational modification peptide motifs.

General examples of types of post-translational modifications include, but are not limited to: (Z)-dehydrobutyryne; 1-chondroitin sulfate-L-aspartic acid ester; 1'-glycosyl-L-tryptophan; 1'-phospho-L-histidine; 1-thioglycine; 2'-(S-L-cysteinyl)-L-histidine; 2'-[3-carboxamido (trimethylammonio)propyl]-L-histidine; 2'-alpha-mannosyl-L-tryptophan; 2-methyl-L-glutamine; 2-oxobutanoic acid; 2-pyrrolidone carboxylic acid; 3'-(1'-L-histidyl)-L-tyrosine; 3'-(8alpha-FAD)-L-histidine; 3'-(S-L-cysteinyl)-L-tyrosine; 3', 3'', 5'-triiodo-L-tyronine; 3'-4'-phospho-L-tyrosine; 3-hydroxy-L-proline; 3'-methyl-L-histidine; 3-methyl-L-lanthionine; 3'-phospho-L-histidine; 4'-(L-tryptophan)-L-tryptophyl quinone; 42 N-cysteinyl-glycosylphosphatidylinositoethanolamine; 43 -(T-L-histidyl)-L-tyrosine; 4-hydroxy-L-arginine; 4-hydroxy-L-lysine; 4-hydroxy-L-proline; 5'-(N6-L-lysine)-L-topaquinone; 5-hydroxy-L-lysine; 5-methyl-L-arginine; alpha-l-microglobulin-Ig alpha complex chromophore; bis-L-cysteinyl bis-L-histidino diiron disulfide; bis-L--cysteinyl-L-N3'-histidino-L-serinyl tetrairon' tetrasulfide; chondroitin sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-serine; D-alanine; D-allo-isoleucine; D-asparagine; dehydroalanine; dehydrotyrosine; dermatan 4-sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-serine; D-glucuronyl-N-glycine; dipyrrolylmethanemethyl-L-cysteine; D-leucine; D-methionine; D-phenylalanine; D-serine; D-tryptophan; glycine amide; glycine oxazolecarboxylic acid; glycine thiazolecarboxylic acid; heme P450-bis-L-cysteine-L-tyrosine; heme-bis-L-cysteine; hemediol-L-aspartyl ester-L-glutamyl ester; hemediol-L-aspartyl ester-L-glutamyl ester-L-methionine sulfonium; heme-L-cysteine; heme-L-histidine; heparan sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-serine; heme P450-bis-L-cysteine-L-lysine; hexakis-L-cysteinyl hexairon hexasulfide; 30 keratan sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-threonine; L-oxoalanine- lactic acid; L phenyllactic acid; 1'-(8alpha-FAD)-L-histidine; L-2'.4',5'-topaquinone; L-3',4'-dihydroxyphenylalanine; L-3'.4'.5'-trihydroxyphenylalanine; L-4'-

- bromophenylalanine; L-6'-bromotryptophan; L-alanine amide; L-alanyl imidazolinone glycine; L-allysine; L-arginine amide; L-asparagine amide; L-aspartic 4-phosphoric anhydride; L-aspartic acid 1-amide; L-beta-methylthioaspartic acid; L-bromohistidine; L-citrulline; L-cysteine amide; L-cysteine glutathione disulfide; L-cysteine methyl disulfide;
- 5 L-cysteine methyl ester; L-cysteine oxazolecarboxylic acid; L-cysteine oxazolinecarboxylic acid; L-cysteine persulfide; L-cysteine sulfenic acid; L-cysteine sulfinic acid; L-cysteine thiazolecarboxylic acid; L-cysteiny l homocitryl molybdenum-heptairon-nonasulfide; L-cysteiny l imidazolinone glycine; L-cysteiny l molybdopterin; L-cysteiny l molybdopterin guanine dinucleotide; L-cystine; L-erythro-beta-
- 10 hydroxyasparagine; L-erythro-beta-hydroxyaspartic acid; L-gamma-carboxyglutamic acid; L-glutamic acid 1-amide; L-glutamic acid 5-methyl ester; L-glutamine amide; L-glutamyl 5-glycerylphosphorylethanolamine; L-histidine amide; L-isoglutamyl-polyglutamic acid; L-isoglutamyl-polyglycine; L-isoleucine amide; L-lanthionine; L-leucine amide; L-lysine amide; L-lysine thiazolecarboxylic acid; L-lysinoalanine; L-methionine amide; L-
- 15 methionine sulfone; L-phenylalanine thiazolecarboxylic acid; L-phenylalanine amide; L-proline amide; L-selenocysteine; L-selenocysteiny l molybdopterin guanine dinucleotide; L-serine amide; L-serine thiazolecarboxylic acid; L-seryl imidazolinone glycine; L-T-bromophenylalanine; L-T-bromophenylalanine; L-threonine amide; L-thyroxine; L-tryptophan amide; L-tryptophyl quinone; L-tyrosine amide; L-valine amide; meso-
- 20 lanthionine; N-(L-glutamyl)-L-tyrosine; N-(L-isoaspartyl)-glycine; N-(L-isoaspartyl)-L-cysteine; N,N,N-trimethyl-L-alanine; N,N-dimethyl-L-proline; N2-acetyl-L-lysine; N2-succinyl-L-tryptophan; N4-(ADP-ribosyl)-L-asparagine; N4-glycosyl-L-asparagine; N4-hydroxymethyl-L-asparagine; N4-methyl-L-asparagine; N5-methyl-L-glutamine; N6- 1 -carboxyethyl-L-lysine; N6-(4-amino hydroxybutyl)-L-lysine; N6-(L-isoglutamyl)-L-
- 25 lysine; N6-(phospho-5'-adenosine)-L-lysine; N6-(phospho-5'-guanosine)-L-lysine; N6,N6,N6-trimethyl-L-lysine; N6,N6-dimethyl-L-lysine; N6-acetyl-L-lysine; N6-biotinyl-L-lysine; N6-carboxy-L-lysine; N6-formyl-L-lysine; N6-glycyl-L-lysine; N6-lipoyl-L-lysine; N6-methyl-L-lysine; N6-methyl-N6-poly(N-methyl-propylamine)-L-lysine; N6-mureinyl-L-lysine; N6-myristoyl-L-lysine; N6-palmitoyl-L-lysine; N6-pyridoxal
- 30 phosphate-L-lysine; N6-pyruvic acid 2-iminyl-L-lysine; N6-retinal-L-lysine; N-acetyl-glycine; N-acetyl-L-glutamine; N-acetyl-L-alanine; N-acetyl-L-aspartic acid; N-acetyl-L-cysteine; N-acetyl-L-glutamic acid; N-acetyl-L-isoleucine; N-acetyl-L-methionine; N-acetyl-L-proline; N-acetyl-L-serine; N-acetyl-L-threonine; N-acetyl-L-

tyrosine; N-acetyl-L-valine; N-alanyl-glycosylphosphatidylinositoethanolamine; N-asparaginyl-glycosylphosphatidylinositoethanolamine; N-aspartyl-glycosylphosphatidylinositoethanolamine; N-formylglycine; N-formyl-L-methionine; N-glycyl-glycosylphosphatidylinositoethanolamine; N-L-glutamyl-poly-L-glutamic acid; N-methylglycine; N-methyl-L-alanine; N-methyl-L-methionine; N-methyl-L-phenylalanine; N-myristoyl-glycine; N-palmitoyl-L-cysteine; N-pyruvic acid 2-iminyl-L-cysteine; N-pyruvic acid 2-iminyl-L-valine; N-seryl-glycosylphosphatidylinositoethanolamine; N-seryl-glycosylphosphatidylcholine; O-(ADP-ribosyl)-L-serine; O-(phospho-5'-adenosine)-L-threonine; O-(phospho-5'-DNA)-L-serine; O-(phospho-5'-DNA)-L-threonine; O-(phospho-5'rRNA)-L-serine; O-(phosphoribosyl dephospho-coenzyme A)-L-serine; O-(sn-1-glycerophosphoryl)-L-serine; O4'-(8alpha-FAD)-L-tyrosine; O4'-(phospho-5'-adenosine)-L-tyrosine; O4'-(phospho-5'-DNA)-L-tyrosine; O4'-(phospho-5'-RNA)-L-tyrosine; O4'-(phospho-5'-uridine)-L-tyrosine; O4-glycosyl-L-hydroxyproline; O4'-glycosyl-L-tyrosine; O4'-sulfo-L-tyrosine; O5-glycosyl-L-hydroxylysine; O-glycosyl-L-serine; O-glycosyl-L-threonine; omega-N-(ADP-ribosyl)-L-arginine; omega-N-omega-N'-dimethyl-L-arginine; omega-N-methyl-L-arginine; omega-N-omega-N-dimethyl-L-arginine; omega-N-phospho-L-arginine; O'octanoyl-L-serine; O-palmitoyl-L-serine; O-palmitoyl-L-threonine; O-phospho-L-serine; O-phospho-L-threonine; O-phosphopantetheine-L-serine; phycoerythrobilin-bis-L-cysteine; phycourobilin-bis-L-cysteine; pyrroloquinoline quinone; pyruvic acid; S hydroxycinnamyl-L-cysteine; S-(2-aminovinyl) methyl-D-cysteine; S-(2-aminovinyl)-D-cysteine; S-(6-FW)-L-cysteine; S-(8alpha-FAD)-L-cysteine; S-(ADP-ribosyl)-L-cysteine; S-(L-isoglutamyl)-L-cysteine; S-12-hydroxyfarnesyl-L-cysteine; S-acetyl-L-cysteine; S-diacylglycerol-L-cysteine; S-diphytanylglycerol diether-L-cysteine; S-farnesyl-L-cysteine; S-geranylgeranyl-L-cysteine; S-glycosyl-L-cysteine; S-glycyl-L-cysteine; S-methyl-L-cysteine; S-nitrosyl-L-cysteine; S-palmitoyl-L-cysteine; S-phospho-L-cysteine; S-phycobiliviolin-L-cysteine; S-phycocyanobilin-L-cysteine; S-phycoerythrobilin-L-cysteine; S-phytochromobilin-L-cysteine; S-selenyl-L-cysteine; S-sulfo-L-cysteine; tetrakis-L-cysteiny diiron disulfide; tetrakis-L-cysteiny iron; tetrakis-L-cysteiny tetrairon tetrasulfide; trans-2,3-cis 4-dihydroxy-L-proline; tris-L-cysteiny triiron tetrasulfide; tris-L-cysteiny triiron trisulfide; tris-L-cysteiny-L-aspartato tetrairon tetrasulfide; tris-L-cysteiny-L-cysteine persulfido-bis-L-glutamato-L-histidino tetrairon disulfide trioxide; tris-L-cysteiny-L-N3'-histidino-

tetrairon tetrasulfide; tris-L-cysteinyl-L-NI'-histidino tetrairon tetrasulfide; and tris-L-cysteinyl-L-serinyl tetrairon tetrasulfide.

Additional examples of PTMs may be found in web sites such as the Delta Mass database based on Krishna, R. G. and F. Wold (1998). Posttranslational Modifications. 5 Proteins - Analysis and Design. R. H. Angeletti. San Diego, Academic Press. 1: 121-206; Methods in Enzymology, 193, J.A. McClosky (ed) (1990), pages 647-660; Methods in Protein Sequence Analysis edited by Kazutomo Imahori and Fumio Sakiyama, Plenum Press, (1993) "Post-translational modifications of proteins" R.G. Krishna and F. Wold pages 167-172; "GlycoSuiteDB: a new curated relational database of glycoprotein glycan 10 structures and their biological sources" Cooper et al. Nucleic Acids Res. 29; 332-335 (2001) "O-GLYCBASE version 4.0: a revised database of O-glycosylated proteins" Gupta et al. Nucleic Acids Research, 27: 370-372 (1999); and "PhosphoBase, a database of phosphorylation sites: release 2.0.", Kreegipuu et al. Nucleic Acids Res 27(1):237-239 (1999) see also, WO 02/21139A2, the disclosure of which is incorporated herein by 15 reference in its entirety.

Tumorigenesis is often accompanied by alterations in the post-translational modifications of proteins. Thus, in another embodiment, the invention provides polypeptides from cancerous cells or tissues that have altered post-translational modifications compared to the post-translational modifications of polypeptides from 20 normal cells or tissues. A number of altered post-translational modifications are known. One common alteration is a change in phosphorylation state, wherein the polypeptide from the cancerous cell or tissue is hyperphosphorylated or hypophosphorylated compared to the polypeptide from a normal tissue, or wherein the polypeptide is phosphorylated on different residues than the polypeptide from a normal cell. Another common alteration is 25 a change in glycosylation state, wherein the polypeptide from the cancerous cell or tissue has more or less glycosylation than the polypeptide from a normal tissue, and/or wherein the polypeptide from the cancerous cell or tissue has a different type of glycosylation than the polypeptide from a noncancerous cell or tissue. Changes in glycosylation may be critical because carbohydrate-protein and carbohydrate-carbohydrate interactions are 30 important in cancer cell progression, dissemination and invasion. See, e.g., Barchi, *Curr. Pharm. Des.* 6: 485-501 (2000), Verma, *Cancer Biochem. Biophys.* 14: 151-162 (1994) and Dennis et al., *Bioessays* 5: 412-421 (1999).

Another post-translational modification that may be altered in cancer cells is prenylation. Prenylation is the covalent attachment of a hydrophobic prenyl group (either farnesyl or geranylgeranyl) to a polypeptide. Prenylation is required for localizing a protein to a cell membrane and is often required for polypeptide function. For instance,
5 the Ras superfamily of GTPase signalling proteins must be prenylated for function in a cell. See, e.g., Prendergast et al., *Semin. Cancer Biol.* 10: 443-452 (2000) and Khwaja et al., *Lancet* 355: 741-744 (2000).

Other post-translation modifications that may be altered in cancer cells include, without limitation, polypeptide methylation, acetylation, arginylation or racemization of
10 amino acid residues. In these cases, the polypeptide from the cancerous cell may exhibit either increased or decreased amounts of the post-translational modification compared to the corresponding polypeptides from noncancerous cells.

Other polypeptide alterations in cancer cells include abnormal polypeptide cleavage of proteins and aberrant protein-protein interactions. Abnormal polypeptide
15 cleavage may be cleavage of a polypeptide in a cancerous cell that does not usually occur in a normal cell, or a lack of cleavage in a cancerous cell, wherein the polypeptide is cleaved in a normal cell. Aberrant protein-protein interactions may be either covalent cross-linking or non-covalent binding between proteins that do not normally bind to each other. Alternatively, in a cancerous cell, a protein may fail to bind to another protein to
20 which it is bound in a noncancerous cell. Alterations in cleavage or in protein-protein interactions may be due to over- or underproduction of a polypeptide in a cancerous cell compared to that in a normal cell, or may be due to alterations in post-translational modifications (see above) of one or more proteins in the cancerous cell. See, e.g., Henschen-Edman, *Ann. N.Y. Acad. Sci.* 936: 580-593 (2001).

25 Alterations in polypeptide post-translational modifications, as well as changes in polypeptide cleavage and protein-protein interactions, may be determined by any method known in the art. For instance, alterations in phosphorylation may be determined by using anti-phosphoserine, anti-phosphothreonine or anti-phosphotyrosine antibodies or by amino acid analysis. Glycosylation alterations may be determined using antibodies specific for
30 different sugar residues, by carbohydrate sequencing, or by alterations in the size of the glycoprotein, which can be determined by, e.g., SDS polyacrylamide gel electrophoresis (PAGE). Other alterations of post-translational modifications, such as prenylation, racemization, methylation, acetylation and arginylation, may be determined by chemical

analysis, protein sequencing, amino acid analysis, or by using antibodies specific for the particular post-translational modifications. Changes in protein-protein interactions and in polypeptide cleavage may be analyzed by any method known in the art including, without limitation, non-denaturing PAGE (for non-covalent protein-protein interactions), SDS
5 PAGE (for covalent protein-protein interactions and protein cleavage), chemical cleavage, protein sequencing or immunoassays.

In another embodiment, the invention provides polypeptides that have been post-translationally modified. In one embodiment, polypeptides may be modified enzymatically or chemically, by addition or removal of a post-translational modification.
10 For example, a polypeptide may be glycosylated or deglycosylated enzymatically. Similarly, polypeptides may be phosphorylated using a purified kinase, such as a MAP kinase (e.g., p38, ERK, or JNK) or a tyrosine kinase (e.g., Src or erbB2). A polypeptide may also be modified through synthetic chemistry. Alternatively, one may isolate the polypeptide of interest from a cell or tissue that expresses the polypeptide with the desired
15 post-translational modification. In another embodiment, a nucleic acid molecule encoding the polypeptide of interest is introduced into a host cell that is capable of post-translationally modifying the encoded polypeptide in the desired fashion. If the polypeptide does not contain a motif for a desired post-translational modification, one may alter the post-translational modification by mutating the nucleic acid sequence of a nucleic
20 acid molecule encoding the polypeptide so that it contains a site for the desired post-translational modification. Amino acid sequences that may be post-translationally modified are known in the art. See, e.g., the programs described above on the website expasy.org of the world wide web. The nucleic acid molecule may also be introduced into a host cell that is capable of post-translationally modifying the encoded polypeptide.
25 Similarly, one may delete sites that are post-translationally modified by either mutating the nucleic acid sequence so that the encoded polypeptide does not contain the post-translational modification motif, or by introducing the native nucleic acid molecule into a host cell that is not capable of post-translationally modifying the encoded polypeptide.

It will be appreciated, as is well known and as noted above, that polypeptides are
30 not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of posttranslational events, including natural processing events and events brought about by human manipulation which do not occur naturally. Circular, branched and branched

circular polypeptides may be synthesized by non-translation natural processes and by entirely synthetic methods, as well. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl group in a polypeptide, or both, by a covalent modification, is common in naturally occurring and synthetic polypeptides and such modifications may be present in polypeptides of the present invention, as well. For instance, the amino terminal residue of polypeptides made in *E. coli*, prior to proteolytic processing, almost invariably will be N-formylmethionine.

Useful post-synthetic (and post-translational) modifications include conjugation to detectable labels, such as fluorophores. A wide variety of amine-reactive and thiol-reactive fluorophore derivatives have been synthesized that react under nondenaturing conditions with N-terminal amino groups and epsilon amino groups of lysine residues, on the one hand, and with free thiol groups of cysteine residues, on the other.

Kits are available commercially that permit conjugation of proteins to a variety of amine-reactive or thiol-reactive fluorophores: Molecular Probes, Inc. (Eugene, OR, USA), *e.g.*, offers kits for conjugating proteins to Alexa Fluor 350, Alexa Fluor 430, Fluorescein-EX, Alexa Fluor 488, Oregon Green 488, Alexa Fluor 532, Alexa Fluor 546, Alexa Fluor 568, Alexa Fluor 594, and Texas Red-X.

A wide variety of other amine-reactive and thiol-reactive fluorophores are available commercially (Molecular Probes, Inc., Eugene, OR, USA), including Alexa Fluor® 350, Alexa Fluor® 488, Alexa Fluor® 532, Alexa Fluor® 546, Alexa Fluor® 568, Alexa Fluor® 594, Alexa Fluor® 647 (monoclonal antibody labeling kits available from Molecular Probes, Inc., Eugene, OR, USA), BODIPY dyes, such as BODIPY 493/503, BODIPY FL, BODIPY R6G, BODIPY 530/550, BODIPY TMR, BODIPY 558/568, BODIPY 558/568, BODIPY 564/570, BODIPY 576/589, BODIPY 581/591, BODIPY TR, BODIPY 630/650, BODIPY 650/665, Cascade Blue, Cascade Yellow, Dansyl, lissamine rhodamine B, Marina Blue, Oregon Green 488, Oregon Green 514, Pacific Blue, rhodamine 6G, rhodamine green, rhodamine red, tetramethylrhodamine, Texas Red (available from Molecular Probes, Inc., Eugene, OR, USA).

The polypeptides of the present invention can also be conjugated to fluorophores, other proteins, and other macromolecules, using bifunctional linking reagents. Common homobifunctional reagents include, *e.g.*, APG, AEDP, BASED, BMB, BMDB, BMH, BMOE, BM[PEO]3, BM[PEO]4, BS3, BSOE, DFDNB, DMA, DMP, DMS, DPDPB,

DSG, DSP (Lomant's Reagent), DSS, DST, DTBP, DTME, DTSSP, EGS, HBVS, Sulfo-BSOCOES, Sulfo-DST, Sulfo-EGS (all available from Pierce, Rockford, IL, USA); common heterobifunctional cross-linkers include ABH, AMAS, ANB-NOS, APDP, ASBA, BMPA, BMPH, BMPS, EDC, EMCA, EMCH, EMCS, KMUA, KMUH, GMBS, LC-SMCC, LC-SPDP, MBS, M2C2H, MPBH, MSA, NHS-ASA, PDPH, PMPI, SADP, SAED, SAND, SANPAH, SASD, SATP, SBAP, SFAD, SIA, SIAB, SMCC, SMPB, SMPT, SPDP, Sulfo-EMCS, Sulfo-GMBS, Sulfo-HSAB, Sulfo-KMUS, Sulfo-LC-SPDP, Sulfo-MBS, Sulfo-NHS-LC-ASA, Sulfo-SADP, Sulfo-SANPAH, Sulfo-SIAB, Sulfo-SMCC, Sulfo-SMPB, Sulfo-LC-SMPT, SVSB, TFCS (all available
10 Pierce, Rockford, IL, USA).

Polypeptides of the present invention, including full length polypeptides, fragments and fusion proteins, can be conjugated, using such cross-linking reagents, to fluorophores that are not amine- or thiol-reactive. Other labels that usefully can be conjugated to polypeptides of the present invention include radioactive labels,
15 echosonographic contrast reagents, and MRI contrast agents.

Polypeptides of the present invention, including full length polypeptides, fragments and fusion proteins, can also usefully be conjugated using cross-linking agents to carrier proteins, such as KLH, bovine thyroglobulin, and even bovine serum albumin (BSA), to increase immunogenicity for raising anti-CSP antibodies.

Polypeptides of the present invention, including full length polypeptides, fragments and fusion proteins, can also usefully be conjugated to polyethylene glycol (PEG); PEGylation increases the serum half life of proteins administered intravenously for replacement therapy. Delgado *et al.*, *Crit. Rev. Ther. Drug Carrier Syst.* 9(3-4): 249-304 (1992); Scott *et al.*, *Curr. Pharm. Des.* 4(6): 423-38 (1998); DeSantis *et al.*, *Curr. Opin. Biotechnol.* 10(4): 324-30 (1999). PEG monomers can be attached to the protein directly or through a linker, with PEGylation using PEG monomers activated with tresyl chloride (2,2,2-trifluoroethanesulphonyl chloride) permitting direct attachment under mild conditions.

Polypeptides of the present invention are also inclusive of analogs of a polypeptide encoded by a nucleic acid molecule according to the instant invention. In a preferred
30 embodiment, this polypeptide is a CSP. In a more preferred embodiment, this polypeptide is derived from a polypeptide having part or all of the amino acid sequence of SEQ ID NO: 96-237. Also preferred is an analog polypeptide comprising one or more

substitutions of non-natural amino acids or non-native inter-residue bonds compared to the naturally occurring polypeptide. In one embodiment, the analog is structurally similar to a CSP, but one or more peptide linkages is replaced by a linkage selected from the group consisting of --CH₂NH--, --CH₂S--, --CH₂-CH₂--, --CH=CH--(cis and trans), --COCH₂--,
5 --CH(OH)CH₂-- and --CH₂SO--. In another embodiment, the analog comprises substitution of one or more amino acids of a CSP with a D-amino acid of the same type or other non-natural amino acid in order to generate more stable peptides. D-amino acids can readily be incorporated during chemical peptide synthesis: peptides assembled from D-amino acids are more resistant to proteolytic attack; incorporation of D-amino acids can
10 also be used to confer specific three-dimensional conformations on the peptide. Other amino acid analogues commonly added during chemical synthesis include ornithine, norleucine, phosphorylated amino acids (typically phosphoserine, phosphothreonine, phosphotyrosine), L-malonyltyrosine, a non-hydrolyzable analog of phosphotyrosine (*see, e.g., Kote et al., Biochem. Biophys. Res. Com.* 209: 817-821 (1995)), and various
15 halogenated phenylalanine derivatives.

Non-natural amino acids can be incorporated during solid phase chemical synthesis or by recombinant techniques, although the former is typically more common. Solid phase chemical synthesis of peptides is well established in the art. Procedures are described, *inter alia*, in Chan *et al.* (eds.), Fmoc Solid Phase Peptide Synthesis: A
20 Practical Approach (Practical Approach Series), Oxford Univ. Press (March 2000); Jones, Amino Acid and Peptide Synthesis (Oxford Chemistry Primers, No 7), Oxford Univ. Press (1992); and Bodanszky, Principles of Peptide Synthesis (Springer Laboratory), Springer Verlag (1993).

Amino acid analogues having detectable labels are also usefully incorporated
25 during synthesis to provide derivatives and analogs. Biotin, for example can be added using biotinoyl-(9-fluorenylmethoxycarbonyl)-L-lysine (Fmoc biocytin) (Molecular Probes, Eugene, OR, USA). Biotin can also be added enzymatically by incorporation into a fusion protein of an *E. coli* BirA substrate peptide. The Fmoc and tBOC derivatives of dabcyL-L-lysine (Molecular Probes, Inc., Eugene, OR, USA) can be used to incorporate
30 the dabcyL chromophore at selected sites in the peptide sequence during synthesis. The aminonaphthalene derivative EDANS, the most common fluorophore for pairing with the dabcyL quencher in fluorescence resonance energy transfer (FRET) systems, can be

introduced during automated synthesis of peptides by using EDANS-FMOC-L-glutamic acid or the corresponding *t*BOC derivative (both from Molecular Probes, Inc., Eugene, OR, USA). Tetramethylrhodamine fluorophores can be incorporated during automated FMOC synthesis of peptides using (FMOC)-TMR-L-lysine (Molecular Probes, Inc. Eugene, OR, USA).

Other useful amino acid analogues that can be incorporated during chemical synthesis include aspartic acid, glutamic acid, lysine, and tyrosine analogues having allyl side-chain protection (Applied Biosystems, Inc., Foster City, CA, USA); the allyl side chain permits synthesis of cyclic, branched-chain, sulfonated, glycosylated, and phosphorylated peptides.

A large number of other FMOC-protected non-natural amino acid analogues capable of incorporation during chemical synthesis are available commercially, including, e.g., Fmoc-2-aminobicyclo[2.2.1]heptane-2-carboxylic acid, Fmoc-3-endo-aminobicyclo[2.2.1]heptane-2-endo-carboxylic acid, Fmoc-3-exo-aminobicyclo[2.2.1]heptane-2-exo-carboxylic acid, Fmoc-3-endo-amino-bicyclo[2.2.1]hept-5-ene-2-endo-carboxylic acid, Fmoc-3-exo-amino-bicyclo[2.2.1]hept-5-ene-2-exo-carboxylic acid, Fmoc-cis-2-amino-1-cyclohexanecarboxylic acid, Fmoc-trans-2-amino-1-cyclohexanecarboxylic acid, Fmoc-1-amino-1-cyclopentanecarboxylic acid, Fmoc-cis-2-amino-1-cyclopentanecarboxylic acid, Fmoc-1-amino-1-cyclopropanecarboxylic acid, Fmoc-D-2-amino-4-(ethylthio)butyric acid, Fmoc-L-2-amino-4-(ethylthio)butyric acid, Fmoc-L-buthionine, Fmoc-S-methyl-L-Cysteine, Fmoc-2-aminobenzoic acid (anthranillic acid), Fmoc-3-aminobenzoic acid, Fmoc-4-aminobenzoic acid, Fmoc-2-aminobenzophenone-2'-carboxylic acid, Fmoc-N-(4-aminobenzoyl)- β -alanine, Fmoc-2-amino-4,5-dimethoxybenzoic acid, Fmoc-4-aminohippuric acid, Fmoc-2-amino-3-hydroxybenzoic acid, Fmoc-2-amino-5-hydroxybenzoic acid, Fmoc-3-amino-4-hydroxybenzoic acid, Fmoc-4-amino-3-hydroxybenzoic acid, Fmoc-4-amino-2-hydroxybenzoic acid, Fmoc-5-amino-2-hydroxybenzoic acid, Fmoc-2-amino-3-methoxybenzoic acid, Fmoc-4-amino-3-methoxybenzoic acid, Fmoc-2-amino-3-methylbenzoic acid, Fmoc-2-amino-5-methylbenzoic acid, Fmoc-2-amino-6-methylbenzoic acid, Fmoc-3-amino-2-methylbenzoic acid, Fmoc-3-amino-4-methylbenzoic acid, Fmoc-4-amino-3-methylbenzoic acid, Fmoc-3-amino-2-naphtoic acid, Fmoc-D,L-3-amino-3-phenylpropionic acid, Fmoc-L-Methyl-dopa, Fmoc-2-amino-4,6-dimethyl-3-

pyridinecarboxylic acid, Fmoc-D,L-amino-2-thiophenacetic acid, Fmoc-4-(carboxymethyl)piperazine, Fmoc-4-carboxypiperazine, Fmoc-4-(carboxymethyl)homopiperazine, Fmoc-4-phenyl-4-piperidinecarboxylic acid, Fmoc-L-1,2,3,4-tetrahydronorharman-3-carboxylic acid, Fmoc-L-thiazolidine-4-carboxylic acid, all
5 available from The Peptide Laboratory (Richmond, CA, USA).

Non-natural residues can also be added biosynthetically by engineering a suppressor tRNA, typically one that recognizes the UAG stop codon, by chemical aminoacylation with the desired unnatural amino acid. Conventional site-directed mutagenesis is used to introduce the chosen stop codon UAG at the site of interest in the
10 protein gene. When the acylated suppressor tRNA and the mutant gene are combined in an *in vitro* transcription/translation system, the unnatural amino acid is incorporated in response to the UAG codon to give a protein containing that amino acid at the specified position. Liu *et al.*, *Proc. Natl Acad. Sci. USA* 96(9): 4780-5 (1999); Wang *et al.*, *Science* 292(5516): 498-500 (2001).

15 *Fusion Proteins*

Another aspect of the present invention relates to the fusion of a polypeptide of the present invention to heterologous polypeptides. In a preferred embodiment, the polypeptide of the present invention is a CSP. In a more preferred embodiment, the polypeptide of the present invention that is fused to a heterologous polypeptide which
20 comprises part or all of the amino acid sequence of SEQ ID NO: 96-237, or is a mutein, homologous polypeptide, analog or derivative thereof. In an even more preferred embodiment, the fusion protein is encoded by a nucleic acid molecule comprising all or part of the nucleic acid sequence of SEQ ID NO: 1-95, or comprises all or part of a nucleic acid sequence that selectively hybridizes or is homologous to a nucleic acid molecule
25 comprising a nucleic acid sequence of SEQ ID NO: 1-95.

The fusion proteins of the present invention will include at least one fragment of a polypeptide of the present invention, which fragment is at least 6, typically at least 8, often at least 15, and usefully at least 16, 17, 18, 19, or 20 amino acids long. The fragment of the polypeptide of the present to be included in the fusion can usefully be at least 25
30 amino acids long, at least 50 amino acids long, and can be at least 75, 100, or even 150 amino acids long. Fusions that include the entirety of a polypeptide of the present invention have particular utility.

The heterologous polypeptide included within the fusion protein of the present invention is at least 6 amino acids in length, often at least 8 amino acids in length, and preferably at least 15, 20, or 25 amino acids in length. Fusions that include larger polypeptides, such as the IgG Fc region, and even entire proteins (such as GFP
5 chromophore-containing proteins) are particularly useful.

As described above in the description of vectors and expression vectors of the present invention, which discussion is incorporated here by reference in its entirety, heterologous polypeptides to be included in the fusion proteins of the present invention can usefully include those designed to facilitate purification and/or visualization of
10 recombinantly-expressed proteins. *See, e.g.*, Ausubel, Chapter 16, (1992), *supra*. Although purification tags can also be incorporated into fusions that are chemically synthesized, chemical synthesis typically provides sufficient purity that further purification by HPLC suffices; however, visualization tags as above described retain their utility even when the protein is produced by chemical synthesis, and when so included
15 render the fusion proteins of the present invention useful as directly detectable markers of the presence of a polypeptide of the invention.

As also discussed above, heterologous polypeptides to be included in the fusion proteins of the present invention can usefully include those that facilitate secretion of recombinantly expressed proteins into the periplasmic space or extracellular milieu for
20 prokaryotic hosts or into the culture medium for eukaryotic cells through incorporation of secretion signals and/or leader sequences. For example, a His⁶ tagged protein can be purified on a Ni affinity column and a GST fusion protein can be purified on a glutathione affinity column. Similarly, a fusion protein comprising the Fc domain of IgG can be purified on a Protein A or Protein G column and a fusion protein comprising an epitope
25 tag such as myc can be purified using an immunoaffinity column containing an anti-c-myc antibody. It is preferable that the epitope tag be separated from the protein encoded by the essential gene by an enzymatic cleavage site that can be cleaved after purification. See also the discussion of nucleic acid molecules encoding fusion proteins that may be expressed on the surface of a cell.

30 Other useful fusion proteins of the present invention include those that permit use of the polypeptide of the present invention as bait in a yeast two-hybrid system. *See* Bartel *et al.* (eds.), The Yeast Two-Hybrid System, Oxford University Press (1997); Zhu *et al.*, Yeast Hybrid Technologies, Eaton Publishing (2000); Fields *et al.*, *Trends Genet.*

10(S): 286-92 (1994); Mendelsohn *et al.*, *Curr. Opin. Biotechnol.* 5(5): 482-6 (1994); Luban *et al.*, *Curr. Opin. Biotechnol.* 6(1): 59-64 (1995); Allen *et al.*, *Trends Biochem. Sci.* 20(12): 511-6 (1995); Drees, *Curr. Opin. Chem. Biol.* 3(1): 64-70 (1999); Topcu *et al.*, *Pharm. Res.* 17(9): 1049-55 (2000); Fashena *et al.*, *Gene* 250(1-2): 1-14 (2000); Colas
5 *et al.*, *Nature* 380, 548-550 (1996); Norman, T. *et al.*, *Science* 285, 591-595 (1999); Fabbri *et al.*, *Oncogene* 18, 4357-4363 (1999); Xu *et al.*, *Proc Natl Acad Sci U S A.* 94, 12473-12478 (1997); Yang, *et al.*, *Nuc. Acids Res.* 23, 1152-1156 (1995); Kolonin *et al.*, *Proc Natl Acad Sci U S A* 95, 14266-14271 (1998); Cohen *et al.*, *Proc Natl Acad Sci U S A* 95, 14272-14277 (1998); Uetz, *et al.* *Nature* 403, 623-627(2000); Ito, *et al.*, *Proc Natl*
10 *Acad Sci U S A* 98, 4569-4574 (2001). Typically, such fusion is to either *E. coli* LexA or yeast GAL4 DNA binding domains. Related bait plasmids are available that express the bait fused to a nuclear localization signal.

Other useful fusion proteins include those that permit display of the encoded polypeptide on the surface of a phage or cell, fusions to intrinsically fluorescent proteins,
15 such as green fluorescent protein (GFP), and fusions to the IgG Fc region, as described above.

The polypeptides of the present invention can also usefully be fused to protein toxins, such as *Pseudomonas* exotoxin A, diphtheria toxin, shiga toxin A, anthrax toxin lethal factor, or ricin, in order to effect ablation of cells that bind or take up the proteins of
20 the present invention.

Fusion partners include, *inter alia*, *myc*, hemagglutinin (HA), GST, immunoglobulins, β -galactosidase, biotin *trpE*, protein A, β -lactamase, α -amylase, maltose binding protein, alcohol dehydrogenase, polyhistidine (for example, six histidine at the amino and/or carboxyl terminus of the polypeptide), *lacZ*, green fluorescent protein
25 (GFP), yeast α mating factor, GAL4 transcription activation or DNA binding domain, luciferase, and serum proteins such as ovalbumin, albumin and the constant domain of IgG. *See, e.g.*, Ausubel (1992), *supra* and Ausubel (1999), *supra*. Fusion proteins may also contain sites for specific enzymatic cleavage, such as a site that is recognized by enzymes such as Factor XIII, trypsin, pepsin, or any other enzyme known in the art.
30 Fusion proteins will typically be made by either recombinant nucleic acid methods, as described above, chemically synthesized using techniques well known in the art (*e.g.*, a Merrifield synthesis), or produced by chemical cross-linking.

Another advantage of fusion proteins is that the epitope tag can be used to bind the fusion protein to a plate or column through an affinity linkage for screening binding proteins or other molecules that bind to the CSP.

As further described below, the polypeptides of the present invention can readily
5 be used as specific immunogens to raise antibodies that specifically recognize polypeptides of the present invention including CSPs and their allelic variants and homologues. The antibodies, in turn, can be used, *inter alia*, specifically to assay for the polypeptides of the present invention, particularly CSPs, *e.g.* by ELISA for detection of protein fluid samples, such as serum, by immunohistochemistry or laser scanning
10 cytometry, for detection of protein in tissue samples, or by flow cytometry, for detection of intracellular protein in cell suspensions, for specific antibody-mediated isolation and/or purification of CSPs, as for example by immunoprecipitation, and for use as specific agonists or antagonists of CSPs.

One may determine whether polypeptides of the present invention including CSPs,
15 mutans, homologous proteins or allelic variants or fusion proteins of the present invention are functional by methods known in the art. For instance, residues that are tolerant of change while retaining function can be identified by altering the polypeptide at known residues using methods known in the art, such as alanine scanning mutagenesis, Cunningham *et al.*, *Science* 244(4908): 1081-5 (1989); transposon linker scanning
20 mutagenesis, Chen *et al.*, *Gene* 263(1-2): 39-48 (2001); combinations of homolog- and alanine-scanning mutagenesis, Jin *et al.*, *J. Mol. Biol.* 226(3): 851-65 (1992); and combinatorial alanine scanning, Weiss *et al.*, *Proc. Natl. Acad. Sci USA* 97(16): 8950-4 (2000), followed by functional assay. Transposon linker scanning kits are available commercially (New England Biolabs, Beverly, MA, USA, catalog. no. E7-102S;
25 EZ::TN™ In-Frame Linker Insertion Kit, catalogue no. EZI04KN, (Epicentre Technologies Corporation, Madison, WI, USA).

Purification of the polypeptides or fusion proteins of the present invention is well known and within the skill of one having ordinary skill in the art. *See, e.g.*, Scopes, Protein Purification, 2d ed. (1987). Purification of recombinantly expressed polypeptides
30 is described above. Purification of chemically-synthesized peptides can readily be effected, *e.g.*, by HPLC.

Accordingly, it is an aspect of the present invention to provide the isolated polypeptides or fusion proteins of the present invention in pure or substantially pure form

in the presence or absence of a stabilizing agent. Stabilizing agents include both proteinaceous and non-proteinaceous material and are well known in the art. Stabilizing agents, such as albumin and polyethylene glycol (PEG) are known and are commercially available.

5 Although high levels of purity are preferred when the isolated polypeptide or fusion protein of the present invention are used as therapeutic agents, such as in vaccines and replacement therapy, the isolated polypeptides of the present invention are also useful at lower purity. For example, partially purified polypeptides of the present invention can be used as immunogens to raise antibodies in laboratory animals.

10 In a preferred embodiment, the purified and substantially purified polypeptides of the present invention are in compositions that lack detectable ampholytes, acrylamide monomers, bis-acrylamide monomers, and polyacrylamide.

 The polypeptides or fusion proteins of the present invention can usefully be attached to a substrate. The substrate can be porous or solid, planar or non-planar; the
15 bond can be covalent or noncovalent. For example, the peptides of the invention may be stabilized by covalent linkage to albumin. See, U.S. Patent No. 5,876,969, the contents of which are hereby incorporated in its entirety.

 The polypeptides or fusion proteins of the present invention can also be usefully bound to a porous substrate, commonly a membrane, typically comprising nitrocellulose,
20 polyvinylidene fluoride (PVDF), or cationically derivatized, hydrophilic PVDF; so bound, the polypeptides or fusion proteins of the present invention can be used to detect and quantify antibodies, *e.g.* in serum, that bind specifically to the immobilized polypeptide or fusion protein of the present invention.

 As another example, the polypeptides or fusion proteins of the present invention
25 can usefully be bound to a substantially nonporous substrate, such as plastic, to detect and quantify antibodies, *e.g.* in serum, that bind specifically to the immobilized protein of the present invention. Such plastics include polymethylacrylic, polyethylene, polypropylene, polyacrylate, polymethylmethacrylate, polyvinylchloride, polytetrafluoroethylene, polystyrene, polycarbonate, polyacetal, polysulfone, celluloseacetate, cellulosenitrate,
30 nitrocellulose, or mixtures thereof; when the assay is performed in a standard microtiter dish, the plastic is typically polystyrene.

 The polypeptides and fusion proteins of the present invention can also be attached to a substrate suitable for use as a surface enhanced laser desorption ionization source; so

attached, the polypeptide or fusion protein of the present invention is useful for binding and then detecting secondary proteins that bind with sufficient affinity or avidity to the surface-bound polypeptide or fusion protein to indicate biologic interaction there between. The polypeptides or fusion proteins of the present invention can also be attached to a
5 substrate suitable for use in surface plasmon resonance detection; so attached, the polypeptide or fusion protein of the present invention is useful for binding and then detecting secondary proteins that bind with sufficient affinity or avidity to the surface-bound polypeptide or fusion protein to indicate biological interaction there between.

Alternative Transcripts

10 In another aspect, the present invention provides splice variants of genes and proteins encoded thereby. The identification of a novel splice variant which encodes an amino acid sequence with a novel region can be targeted for the generation of reagents for use in detection and/or treatment of cancer. The novel amino acid sequence may lead to a unique protein structure, protein subcellular localization, biochemical processing or
15 function of the splice variant. This information can be used to directly or indirectly facilitate the generation of additional or novel therapeutics or diagnostics. The nucleotide sequence in this novel splice variant can be used as a nucleic acid probe for the diagnosis and/or treatment of cancer.

Specifically, the newly identified sequences may enable the production of new
20 antibodies or compounds directed against the novel region for use as a therapeutic or diagnostic. Alternatively, the newly identified sequences may alter the biochemical or biological properties of the encoded protein in such a way as to enable the generation of improved or different therapeutics targeting this protein.

Antibodies

25 In another aspect, the invention provides antibodies, including fragments and derivatives thereof, that bind specifically to polypeptides encoded by the nucleic acid molecules of the invention. In a preferred embodiment, the antibodies are specific for a polypeptide that is a CSP, or a fragment, mutein, derivative, analog or fusion protein thereof. In a more preferred embodiment, the antibodies are specific for a polypeptide that
30 comprises SEQ ID NO: 96-237, or a fragment, mutein, derivative, analog or fusion protein thereof.

The antibodies of the present invention can be specific for linear epitopes, discontinuous epitopes, or conformational epitopes of such proteins or protein fragments, either as present on the protein in its native conformation or, in some cases, as present on the proteins as denatured, as, *e.g.*, by solubilization in SDS. New epitopes may also be
5 due to a difference in post translational modifications (PTMs) in disease versus normal tissue. For example, a particular site on a CSP may be glycosylated in cancerous cells, but not glycosylated in normal cells or vice versa. In addition, alternative splice forms of a CSP may be indicative of cancer. Differential degradation of the C or N-terminus of a CSP may also be a marker or target for anticancer therapy. For example, a CSP may be
10 N-terminal degraded in cancer cells exposing new epitopes to antibodies which may selectively bind for diagnostic or therapeutic uses.

As is well known in the art, the degree to which an antibody can discriminate among molecular species in a mixture will depend, in part, upon the conformational relatedness of the species in the mixture; typically, the antibodies of the present invention
15 will discriminate over adventitious binding to non-CSP polypeptides by at least two-fold, more typically by at least 5-fold, typically by more than 10-fold, 25-fold, 50-fold, 75-fold, and often by more than 100-fold, and on occasion by more than 500-fold or 1000-fold. When used to detect the proteins or protein fragments of the present invention, the antibody of the present invention is sufficiently specific when it can be used to determine
20 the presence of the polypeptide of the present invention in samples derived from human colon.

Typically, the affinity or avidity of an antibody (or antibody multimer, as in the case of an IgM pentamer) of the present invention for a protein or protein fragment of the present invention will be at least about 1×10^{-6} molar (M), typically at least about 5×10^{-7}
25 M, 1×10^{-7} M, with affinities and avidities of at least 1×10^{-8} M, 5×10^{-9} M, 1×10^{-10} M and up to 1×10^{-13} M proving especially useful.

The antibodies of the present invention can be naturally occurring forms, such as IgG, IgM, IgD, IgE, IgY, and IgA, from any avian, reptilian, or mammalian species.

Human antibodies can, but will infrequently, be drawn directly from human donors
30 or human cells. In such case, antibodies to the polypeptides of the present invention will typically have resulted from fortuitous immunization, such as autoimmune immunization, with the polypeptide of the present invention. Such antibodies will typically, but will not

invariably, be polyclonal. In addition, individual polyclonal antibodies may be isolated and cloned to generate monoclonals.

Human antibodies are more frequently obtained using transgenic animals that express human immunoglobulin genes, which transgenic animals can be affirmatively immunized with the protein immunogen of the present invention. Human Ig-transgenic mice capable of producing human antibodies and methods of producing human antibodies therefrom upon specific immunization are described, *inter alia*, in U.S. Patent Nos. 6,162,963; 6,150,584; 6,114,598; 6,075,181; 5,939,598; 5,877,397; 5,874,299; 5,814,318; 5,789,650; 5,770,429; 5,661,016; 5,633,425; 5,625,126; 5,569,825; 5,545,807; 5,545,806, and 5,591,669, the disclosures of which are incorporated herein by reference in their entireties. Such antibodies are typically monoclonal, and are typically produced using techniques developed for production of murine antibodies.

Human antibodies are particularly useful, and often preferred, when the antibodies of the present invention are to be administered to human beings as *in vivo* diagnostic or therapeutic agents, since recipient immune response to the administered antibody will often be substantially less than that occasioned by administration of an antibody derived from another species, such as mouse.

IgG, IgM, IgD, IgE, IgY, and IgA antibodies of the present invention are also usefully obtained from other species, including mammals such as rodents (typically mouse, but also rat, guinea pig, and hamster), lagomorphs (typically rabbits), and also larger mammals, such as sheep, goats, cows, and horses; or egg laying birds or reptiles such as chickens or alligators. In such cases, as with the transgenic human-antibody-producing non-human mammals, fortuitous immunization is not required, and the non-human mammal is typically affirmatively immunized, according to standard immunization protocols, with the polypeptide of the present invention. One form of avian antibodies may be generated using techniques described in WO 00/29444, published 25 May 2000, which is herein incorporated by reference in its entirety.

As discussed above, virtually all fragments of 8 or more contiguous amino acids of a polypeptide of the present invention can be used effectively as immunogens when conjugated to a carrier, typically a protein such as bovine thyroglobulin, keyhole limpet hemocyanin, or bovine serum albumin, conveniently using a bifunctional linker such as - those described elsewhere above, which discussion is incorporated by reference here.

Immunogenicity can also be conferred by fusion of the polypeptides of the present invention to other moieties. For example, polypeptides of the present invention can be produced by solid phase synthesis on a branched polylysine core matrix; these multiple antigenic peptides (MAPs) provide high purity, increased avidity, accurate chemical definition and improved safety in vaccine development. Tam *et al.*, *Proc. Natl. Acad. Sci. USA* 85: 5409-5413 (1988); Posnett *et al.*, *J. Biol. Chem.* 263: 1719-1725 (1988).

Protocols for immunizing non-human mammals or avian species are well-established in the art. See Harlow *et al.* (eds.), Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory (1998); Coligan *et al.* (eds.), Current Protocols in Immunology, John Wiley & Sons, Inc. (2001); Zola, Monoclonal Antibodies: Preparation and Use of Monoclonal Antibodies and Engineered Antibody Derivatives (Basics: From Background to Bench), Springer Verlag (2000); Gross M, Speck *J.Dtsch. Tierarztl. Wochenschr.* 103: 417-422 (1996). Immunization protocols often include multiple immunizations, either with or without adjuvants such as Freund's complete adjuvant and Freund's incomplete adjuvant, and may include naked DNA immunization. Moss, *Semin. Immunol.* 2: 317-327 (1990).

Antibodies from non-human mammals and avian species can be polyclonal or monoclonal, with polyclonal antibodies having certain advantages in immunohistochemical detection of the polypeptides of the present invention and monoclonal antibodies having advantages in identifying and distinguishing particular epitopes of the polypeptides of the present invention. Antibodies from avian species may have particular advantage in detection of the polypeptides of the present invention, in human serum or tissues. Vikinge *et al.*, *Biosens. Bioelectron.* 13: 1257-1262 (1998). Following immunization, the antibodies of the present invention can be obtained using any art-accepted technique. Such techniques are well known in the art and are described in detail in references such as Coligan, *supra*; Zola, *supra*; Howard *et al.* (eds.), Basic Methods in Antibody Production and Characterization, CRC Press (2000); Harlow, *supra*; Davis (ed.), Monoclonal Antibody Protocols, Vol. 45, Humana Press (1995); Delves (ed.), Antibody Production: Essential Techniques, John Wiley & Son Ltd (1997); and Kenney, Antibody Solution: An Antibody Methods Manual, Chapman & Hall (1997).

Briefly, such techniques include, *inter alia*, production of monoclonal antibodies by hybridomas and expression of antibodies or fragments or derivatives thereof from host cells engineered to express immunoglobulin genes or fragments thereof. These two

methods of production are not mutually exclusive: genes encoding antibodies specific for the polypeptides of the present invention can be cloned from hybridomas and thereafter expressed in other host cells. Nor need the two necessarily be performed together: *e.g.*, genes encoding antibodies specific for the polypeptides of the present invention can be
5 cloned directly from B cells known to be specific for the desired protein, as further described in U.S. Patent No. 5,627,052, the disclosure of which is incorporated herein by reference in its entirety, or from antibody-displaying phage.

Recombinant expression in host cells is particularly useful when fragments or derivatives of the antibodies of the present invention are desired.

10 Host cells for recombinant antibody production of whole antibodies, antibody fragments, or antibody derivatives can be prokaryotic or eukaryotic.

Prokaryotic hosts are particularly useful for producing phage displayed antibodies of the present invention.

The technology of phage-displayed antibodies, in which antibody variable region
15 fragments are fused, for example, to the gene III protein (pIII) or gene VIII protein (pVIII) for display on the surface of filamentous phage, such as M13, is by now well-established. *See, e.g.*, Sidhu, *Curr. Opin. Biotechnol.* 11(6): 610-6 (2000); Griffiths *et al.*, *Curr. Opin. Biotechnol.* 9(1): 102-8 (1998); Hoogenboom *et al.*, *Immunotechnology*, 4(1): 1-20 (1998); Rader *et al.*, *Current Opinion in Biotechnology* 8: 503-508 (1997); Aujame *et al.*, *Human*
20 *Antibodies* 8: 155-168 (1997); Hoogenboom, *Trends in Biotechnol.* 15: 62-70 (1997); de Kruif *et al.*, 17: 453-455 (1996); Barbas *et al.*, *Trends in Biotechnol.* 14: 230-234 (1996); Winter *et al.*, *Ann. Rev. Immunol.* 433-455 (1994). Techniques and protocols required to generate, propagate, screen (pan), and use the antibody fragments from such libraries have recently been compiled. *See, e.g.*, Barbas (2001), *supra*; Kay, *supra*; and Abelson, *supra*.

25 Typically, phage-displayed antibody fragments are scFv fragments or Fab fragments; when desired, full length antibodies can be produced by cloning the variable regions from the displaying phage into a complete antibody and expressing the full length antibody in a further prokaryotic or a eukaryotic host cell. Eukaryotic cells are also useful for expression of the antibodies, antibody fragments, and antibody derivatives of the
30 present invention. For example, antibody fragments of the present invention can be produced in *Pichia pastoris* and in *Saccharomyces cerevisiae*. *See, e.g.*, Takahashi *et al.*, *Biosci. Biotechnol. Biochem.* 64(10): 2138-44 (2000); Freyre *et al.*, *J. Biotechnol.* 76(2-3):1 57-63 (2000); Fischer *et al.*, *Biotechnol. Appl. Biochem.* 30 (Pt 2): 117-20

(1999); Pennell *et al.*, *Res. Immunol.* 149(6): 599-603 (1998); Eldin *et al.*, *J. Immunol. Methods.* 201(1): 67-75 (1997);, Frenken *et al.*, *Res. Immunol.* 149(6): 589-99 (1998); and Shusta *et al.*, *Nature Biotechnol.* 16(8): 773-7 (1998).

Antibodies, including antibody fragments and derivatives, of the present invention
5 can also be produced in insect cells. *See, e.g.*, Li *et al.*, *Protein Expr. Purif.* 21(1): 121-8 (2001); Ailor *et al.*, *Biotechnol. Bioeng.* 58(2-3): 196-203 (1998); Hsu *et al.*, *Biotechnol. Prog.* 13(1): 96-104 (1997); Edelman *et al.*, *Immunology* 91(1): 13-9 (1997); and Nesbit *et al.*, *J. Immunol. Methods* 151(1-2): 201-8 (1992).

Antibodies and fragments and derivatives thereof of the present invention can also
10 be produced in plant cells, particularly maize or tobacco, Giddings *et al.*, *Nature Biotechnol.* 18(11): 1151-5 (2000); Gavilondo *et al.*, *Biotechniques* 29(1): 128-38 (2000); Fischer *et al.*, *J. Biol. Regul. Homeost. Agents* 14(2): 83-92 (2000); Fischer *et al.*, *Biotechnol. Appl. Biochem.* 30 (Pt 2): 113-6 (1999); Fischer *et al.*, *Biol. Chem.* 380(7-8): 825-39 (1999); Russell, *Curr. Top. Microbiol. Immunol.* 240: 119-38 (1999); and Ma *et al.*, *Plant Physiol.* 109(2): 341-6 (1995).

Antibodies, including antibody fragments and derivatives, of the present invention can also be produced in transgenic, non-human, mammalian milk. *See, e.g.* Pollock *et al.*, *J. Immunol Methods.* 231: 147-57 (1999); Young *et al.*, *Res. Immunol.* 149: 609-10 (1998); and Limonta *et al.*, *Immunotechnology* 1: 107-13 (1995).

20 Mammalian cells useful for recombinant expression of antibodies, antibody fragments, and antibody derivatives of the present invention include CHO cells, COS cells, 293 cells, and myeloma cells. Verma *et al.*, *J. Immunol. Methods* 216(1-2):165-81 (1998) review and compare bacterial, yeast, insect and mammalian expression systems for expression of antibodies. Antibodies of the present invention can also be prepared by cell
25 free translation, as further described in Merk *et al.*, *J. Biochem. (Tokyo)* 125(2): 328-33 (1999) and Ryabova *et al.*, *Nature Biotechnol.* 15(1): 79-84 (1997), and in the milk of transgenic animals, as further described in Pollock *et al.*, *J. Immunol. Methods* 231(1-2): 147-57 (1999).

The invention further provides antibody fragments that bind specifically to one or
30 more of the polypeptides of the present invention or to one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited by one or more of the polypeptides of the present invention or one or more of the polypeptides encoded by the isolated nucleic acid

molecules of the present invention. Among such useful fragments are Fab, Fab', Fv, F(ab)'₂, and single-chain Fv (scFv) fragments. Other useful fragments are described in Hudson, *Curr. Opin. Biotechnol.* 9(4): 395-402 (1998).

5 The present invention also relates to antibody derivatives that bind specifically to one or more of the polypeptides of the present invention, to one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited by one or more of the polypeptides of the present invention or one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention.

10 Among such useful derivatives are chimeric, primatized, and humanized antibodies; such derivatives are less immunogenic in human beings, and thus are more suitable for *in vivo* administration, than are unmodified antibodies from non-human mammalian species. Another useful method is PEGylation to increase the serum half life of the antibodies.

15 Chimeric antibodies typically include heavy and/or light chain variable regions (including both CDR and framework residues) of immunoglobulins of one species, typically mouse, fused to constant regions of another species, typically human. *See, e.g., Morrison et al., Proc. Natl. Acad. Sci USA* 81(21): 6851-5 (1984); Sharon *et al., Nature* 309(5966): 364-7 (1984); Takeda *et al., Nature* 314(6010): 452-4 (1985); and U.S. Patent
20 No. 5,807,715 the disclosure of which is incorporated herein by reference in its entirety. Primatized and humanized antibodies typically include heavy and/or light chain CDRs from a murine antibody grafted into a non-human primate or human antibody V region framework, usually further comprising a human constant region, Riechmann *et al., Nature* 332(6162): 323-7 (1988); Co *et al., Nature* 351(6326): 501-2 (1991); and U.S. Patent Nos.
25 6,054,297; 5,821,337; 5,770,196; 5,766,886; 5,821,123; 5,869,619; 6,180,377; 6,013,256; 5,693,761; and 6,180,370, the disclosures of which are incorporated herein by reference in their entireties. Other useful antibody derivatives of the invention include heteromeric antibody complexes and antibody fusions, such as diabodies (bispecific antibodies), single-chain diabodies, and intrabodies.

30 It is contemplated that the nucleic acids encoding the antibodies of the present invention can be operably joined to other nucleic acids forming a recombinant vector for cloning or for expression of the antibodies of the invention. Accordingly, the present invention includes any recombinant vector containing the coding sequences, or part

thereof, whether for eukaryotic transduction, transfection or gene therapy. Such vectors may be prepared using conventional molecular biology techniques, known to those with skill in the art, and would comprise DNA encoding sequences for the immunoglobulin V-regions including framework and CDRs or parts thereof, and a suitable promoter either
5 with or without a signal sequence for intracellular transport. Such vectors may be transduced or transfected into eukaryotic cells or used for gene therapy (Marasco et al., *Proc. Natl. Acad. Sci. (USA)* 90: 7889-7893 (1993); Duan et al., *Proc. Natl. Acad. Sci. (USA)* 91: 5075-5079 (1994), by conventional techniques, known to those with skill in the art.

10 The antibodies of the present invention, including fragments and derivatives thereof, can usefully be labeled. It is, therefore, another aspect of the present invention to provide labeled antibodies that bind specifically to one or more of the polypeptides of the present invention, to one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited
15 by one or more of the polypeptides of the present invention or one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention. The choice of label depends, in part, upon the desired use.

For example, when the antibodies of the present invention are used for immunohistochemical staining of tissue samples, the label can usefully be an enzyme that
20 catalyzes production and local deposition of a detectable product. Enzymes typically conjugated to antibodies to permit their immunohistochemical visualization are well known, and include alkaline phosphatase, β -galactosidase, glucose oxidase, horseradish peroxidase (HRP), and urease. Typical substrates for production and deposition of visually detectable products include o-nitrophenyl-beta-D-galactopyranoside (ONPG);
25 o-phenylenediamine dihydrochloride (OPD); p-nitrophenyl phosphate (PNPP); p-nitrophenyl-beta-D-galactopyranoside (PNPG); 3',3'-diaminobenzidine (DAB); 3-amino-9-ethylcarbazole (AEC); 4-chloro-1-naphthol (CN);
5-bromo-4-chloro-3-indolyl-phosphate (BCIP); ABTS®; BluoGal; iodonitrotetrazolium (INT); nitroblue tetrazolium chloride (NBT); phenazine methosulfate (PMS);
30 phenolphthalein monophosphate (PMP); tetramethyl benzidine (TMB); tetranitroblue tetrazolium (TNBT); X-Gal; X-Gluc; and X-Glucoside.

Other substrates can be used to produce products for local deposition that are luminescent. For example, in the presence of hydrogen peroxide (H_2O_2), horseradish

peroxidase (HRP) can catalyze the oxidation of cyclic diacylhydrazides, such as luminol. Immediately following the oxidation, the luminol is in an excited state (intermediate reaction product), which decays to the ground state by emitting light. Strong enhancement of the light emission is produced by enhancers, such as phenolic compounds. Advantages
5 include high sensitivity, high resolution, and rapid detection without radioactivity and requiring only small amounts of antibody. *See, e.g., Thorpe et al., Methods Enzymol.* 133: 331-53 (1986); Kricka *et al., J. Immunoassay* 17(1): 67-83 (1996); and Lundqvist *et al., J. Biolumin. Chemilumin.* 10(6): 353-9 (1995). Kits for such enhanced chemiluminescent detection (ECL) are available commercially. The antibodies can also be labeled using
10 colloidal gold.

As another example, when the antibodies of the present invention are used, *e.g.,* for flow cytometric detection, for scanning laser cytometric detection, or for fluorescent immunoassay, they can usefully be labeled with fluorophores. There are a wide variety of fluorophore labels that can usefully be attached to the antibodies of the present invention.
15 For flow cytometric applications, both for extracellular detection and for intracellular detection, common useful fluorophores can be fluorescein isothiocyanate (FITC), allophycocyanin (APC), R-phycoerythrin (PE), peridinin chlorophyll protein (PerCP), Texas Red, Cy3, Cy5, fluorescence resonance energy tandem fluorophores such as PerCP-Cy5.5, PE-Cy5, PE-Cy5.5, PE-Cy7, PE-Texas Red, and APC-Cy7.

Other fluorophores include, *inter alia*, Alexa Fluor® 350, Alexa Fluor® 488, Alexa Fluor® 532, Alexa Fluor® 546, Alexa Fluor® 568, Alexa Fluor® 594, Alexa Fluor® 647 (monoclonal antibody labeling kits available from Molecular Probes, Inc., Eugene, OR, USA), BODIPY dyes, such as BODIPY 493/503, BODIPY FL, BODIPY R6G, BODIPY 530/550, BODIPY TMR, BODIPY 558/568, BODIPY 558/568, BODIPY
25 564/570, BODIPY 576/589, BODIPY 581/591, BODIPY TR, BODIPY 630/650, BODIPY 650/665, Cascade Blue, Cascade Yellow, Dansyl, lissamine rhodamine B, Marina Blue, Oregon Green 488, Oregon Green 514, Pacific Blue, rhodamine 6G, rhodamine green, rhodamine red, tetramethylrhodamine, Texas Red (available from Molecular Probes, Inc., Eugene, OR, USA), and Cy2, Cy3, Cy3.5, Cy5, Cy5.5, Cy7, all of
30 which are also useful for fluorescently labeling the antibodies of the present invention. For secondary detection using labeled avidin, streptavidin, captavidin or neutravidin, the antibodies of the present invention can usefully be labeled with biotin.

When the antibodies of the present invention are used, e.g., for western blotting applications, they can usefully be labeled with radioisotopes, such as ^{33}P , ^{32}P , ^{35}S , ^3H , and ^{125}I . As another example, when the antibodies of the present invention are used for radioimmunotherapy, the label can usefully be ^{228}Th , ^{227}Ac , ^{225}Ac , ^{223}Ra , ^{213}Bi , ^{212}Pb , ^{212}Bi , ^{211}At , ^{203}Pb , ^{194}Os , ^{188}Re , ^{186}Re , ^{153}Sm , ^{149}Tb , ^{131}I , ^{125}I , ^{111}In , ^{105}Rh , $^{99\text{m}}\text{Tc}$, ^{97}Ru , ^{90}Y , ^{90}Sr , ^{88}Y , ^{72}Se , ^{67}Cu , or ^{47}Sc .

As another example, when the antibodies of the present invention are to be used for *in vivo* diagnostic use, they can be rendered detectable by conjugation to MRI contrast agents, such as gadolinium diethylenetriaminepentaacetic acid (DTPA), Lauffer *et al.*,
10 *Radiology* 207(2): 529-38 (1998), or by radioisotopic labeling.

As would be understood, use of the labels described above is not restricted to the application as for which they were mentioned.

The antibodies of the present invention, including fragments and derivatives thereof, can also be conjugated to toxins, in order to target the toxin's ablative action to
15 cells that display and/or express the polypeptides of the present invention. Commonly, the antibody in such immunotoxins is conjugated to *Pseudomonas* exotoxin A, diphtheria toxin, shiga toxin A, anthrax toxin lethal factor, or ricin. See Hall (ed.), Immunotoxin Methods and Protocols (Methods in Molecular Biology, vol. 166), Humana Press (2000); and Frankel *et al.* (eds.), Clinical Applications of Immunotoxins, Springer-Verlag (1998).

20 The antibodies of the present invention can usefully be attached to a substrate, and it is, therefore, another aspect of the invention to provide antibodies that bind specifically to one or more of the polypeptides of the present invention, to one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited by one or more of the polypeptides of
25 the present invention or one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, attached to a substrate. Substrates can be porous or nonporous, planar or nonplanar. For example, the antibodies of the present invention can usefully be conjugated to filtration media, such as NHS-activated Sepharose or CNBr-activated Sepharose for purposes of immunoaffinity chromatography. For example, the
30 antibodies of the present invention can usefully be attached to paramagnetic microspheres, typically by biotin-streptavidin interaction, which microsphere can then be used for isolation of cells that express or display the polypeptides of the present invention. As

another example, the antibodies of the present invention can usefully be attached to the surface of a microtiter plate for ELISA.

As noted above, the antibodies of the present invention can be produced in prokaryotic and eukaryotic cells. It is, therefore, another aspect of the present invention to
5 provide cells that express the antibodies of the present invention, including hybridoma cells, B cells, plasma cells, and host cells recombinantly modified to express the antibodies of the present invention.

In yet a further aspect, the present invention provides aptamers evolved to bind specifically to one or more of the CSPs of the present invention or to polypeptides
10 encoded by the CSNAs of the invention.

In sum, one of skill in the art, provided with the teachings of this invention, has available a variety of methods which may be used to alter the biological properties of the antibodies of this invention including methods which would increase or decrease the stability or half-life, immunogenicity, toxicity, affinity or yield of a given antibody
15 molecule, or to alter it in any other way that may render it more suitable for a particular application.

Transgenic Animals and Cells

In another aspect, the invention provides transgenic cells and non-human organisms comprising nucleic acid molecules of the invention. In a preferred
20 embodiment, the transgenic cells and non-human organisms comprise a nucleic acid molecule encoding a CSP. In a preferred embodiment, the CSP comprises an amino acid sequence selected from SEQ ID NO: 96-237, or a fragment, mutein, homologous protein or allelic variant thereof. In another preferred embodiment, the transgenic cells and non-human organism comprise a CSNA of the invention, preferably a CSNA comprising a
25 nucleotide sequence selected from the group consisting of SEQ ID NO: 1-95, or a part, substantially similar nucleic acid molecule, allelic variant or hybridizing nucleic acid molecule thereof.

In another embodiment, the transgenic cells and non-human organisms have a targeted disruption or replacement of the endogenous orthologue of the human CSG. The
30 transgenic cells can be embryonic stem cells or somatic cells. The transgenic non-human organisms can be chimeric, nonchimeric heterozygotes, and nonchimeric homozygotes. Methods of producing transgenic animals are well known in the art. *See, e.g., Hogan et*

al., Manipulating the Mouse Embryo: A Laboratory Manual, 2d ed., Cold Spring Harbor Press (1999); Jackson *et al.*, Mouse Genetics and Transgenics: A Practical Approach, Oxford University Press (2000); and Pinkert, Transgenic Animal Technology: A Laboratory Handbook, Academic Press (1999).

5 Any technique known in the art may be used to introduce a nucleic acid molecule of the invention into an animal to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection. (*see, e.g.*, Paterson *et al.*, *Appl. Microbiol. Biotechnol.* 40: 691-698 (1994); Carver *et al.*, *Biotechnology* 11: 1263-1270 (1993); Wright *et al.*, *Biotechnology* 9: 830-834 (1991); and U.S. Patent No. 10 4,873,191, herein incorporated by reference in its entirety); retrovirus-mediated gene transfer into germ lines, blastocysts or embryos (*see, e.g.*, Van der Putten *et al.*, *Proc. Natl. Acad. Sci., USA* 82: 6148-6152 (1985)); gene targeting in embryonic stem cells (*see, e.g.*, Thompson *et al.*, *Cell* 56: 313-321 (1989)); electroporation of cells or embryos (*see, e.g.*, Lo, 1983, *Mol. Cell. Biol.* 3: 1803-1814 (1983)); introduction using a gene gun (*see, e.g.*, Ulmer *et al.*, *Science* 259: 1745-49 (1993); introducing nucleic acid constructs into 15 embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-mediated gene transfer (*see, e.g.*, Lavitrano *et al.*, *Cell* 57: 717-723 (1989)).

Other techniques include, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (*see, e.g.*, 20 Campell *et al.*, *Nature* 380: 64-66 (1996); Wilmut *et al.*, *Nature* 385: 810-813 (1997)). The present invention provides for transgenic animals that carry the transgene (*i.e.*, a nucleic acid molecule of the invention) in all their cells, as well as animals which carry the transgene in some, but not all their cells, *i.e.* *e.*, mosaic animals or chimeric animals.

The transgene may be integrated as a single transgene or as multiple copies, such 25 as in concatamers, *e. g.*, head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, *e.g.*, the teaching of Lasko *et al. et al.*, *Proc. Natl. Acad. Sci. USA* 89: 6232- 6236 (1992). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

30 Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the

transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and reverse transcriptase-PCR (RT-PCR). Samples of transgenic gene-expressing tissue may also be evaluated

- 5 immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than
10 one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA
15 analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of
20 the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

Methods for creating a transgenic animal with a disruption of a targeted gene are also well known in the art. In general, a vector is designed to comprise some nucleotide
25 sequences homologous to the endogenous targeted gene. The vector is introduced into a cell so that it may integrate, via homologous recombination with chromosomal sequences, into the endogenous gene, thereby disrupting the function of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type. *See, e.g., Gu et al., Science* 265: 103-106
30 (1994). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. *See, e.g., Smithies et al., Nature* 317: 230-234 (1985); Thomas et al., *Cell* 51: 503-512 (1987); Thompson et al., *Cell* 5: 313-321 (1989).

In one embodiment, a mutant, non-functional nucleic acid molecule of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous nucleic acid sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable
5 marker, to transfect cells that express polypeptides of the invention in vivo. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications
10 to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene. *See, e.g., Thomas, supra* and *Thompson, supra*. However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors that will be apparent to those of skill in the art.

15 In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (*e.g.*, knockouts) are administered to a patient in vivo. Such cells may be obtained from an animal or patient or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells,
20 blood cells (*e.g.*, lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered in vitro using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, *e.g.*, by transduction (using viral vectors, and preferably
25 vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc.

The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve
30 expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, *e.g.*, in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, *e.g.*, genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. *See, e.g.*, U.S. Patent Nos. 5,399,349 and 5,460,959, each of which is
5 incorporated by reference herein in its entirety.

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of
10 components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with
15 aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

Computer Readable Means

A further aspect of the invention is a computer readable means for storing the nucleic acid and amino acid sequences of the instant invention. In a preferred
20 embodiment, the invention provides a computer readable means for storing SEQ ID NO: 96-237 and SEQ ID NO: 1-95 as described herein, as the complete set of sequences or in any combination. The records of the computer readable means can be accessed for reading and display and for interface with a computer system for the application of programs allowing for the location of data upon a query for data meeting certain criteria,
25 the comparison of sequences, the alignment or ordering of sequences meeting a set of criteria, and the like.

The nucleic acid and amino acid sequences of the invention are particularly useful as components in databases useful for search analyses as well as in sequence analysis algorithms. As used herein, the terms "nucleic acid sequences of the invention" and
30 "amino acid sequences of the invention" mean any detectable chemical or physical characteristic of a polynucleotide or polypeptide of the invention that is or may be reduced to or stored in a computer readable form. These include, without limitation,

chromatographic scan data or peak data, photographic data or scan data therefrom, and mass spectrographic data.

This invention provides computer readable media having stored thereon sequences of the invention. A computer readable medium may comprise one or more of the
5 following: a nucleic acid sequence comprising a sequence of a nucleic acid sequence of the invention; an amino acid sequence comprising an amino acid sequence of the invention; a set of nucleic acid sequences wherein at least one of said sequences comprises the sequence of a nucleic acid sequence of the invention; a set of amino acid sequences wherein at least one of said sequences comprises the sequence of an amino acid sequence
10 of the invention; a data set representing a nucleic acid sequence comprising the sequence of one or more nucleic acid sequences of the invention; a data set representing a nucleic acid sequence encoding an amino acid sequence comprising the sequence of an amino acid sequence of the invention; a set of nucleic acid sequences wherein at least one of said sequences comprises the sequence of a nucleic acid sequence of the invention; a set of
15 amino acid sequences wherein at least one of said sequences comprises the sequence of an amino acid sequence of the invention; a data set representing a nucleic acid sequence comprising the sequence of a nucleic acid sequence of the invention; a data set representing a nucleic acid sequence encoding an amino acid sequence comprising the sequence of an amino acid sequence of the invention. The computer readable medium can
20 be any composition of matter used to store information or data, including, for example, commercially available floppy disks, tapes, hard drives, compact disks, and video disks.

Also provided by the invention are methods for the analysis of character sequences, particularly genetic sequences. Preferred methods of sequence analysis include, for example, methods of sequence homology analysis, such as identity and
25 similarity analysis, RNA structure analysis, sequence assembly, cladistic analysis, sequence motif analysis, open reading frame determination, nucleic acid base calling, and sequencing chromatogram peak analysis.

A computer-based method is provided for performing nucleic acid sequence identity or similarity identification. This method comprises the steps of providing a
30 nucleic acid sequence comprising the sequence of a nucleic acid of the invention in a computer readable medium; and comparing said nucleic acid sequence to at least one nucleic acid or amino acid sequence to identify sequence identity or similarity.

A computer-based method is also provided for performing amino acid homology identification, said method comprising the steps of: providing an amino acid sequence comprising the sequence of an amino acid of the invention in a computer readable medium; and comparing said amino acid sequence to at least one nucleic acid or an amino acid sequence to identify homology.

A computer-based method is still further provided for assembly of overlapping nucleic acid sequences into a single nucleic acid sequence, said method comprising the steps of: providing a first nucleic acid sequence comprising the sequence of a nucleic acid of the invention in a computer readable medium; and screening for at least one overlapping region between said first nucleic acid sequence and a second nucleic acid sequence. In addition, the invention includes a method of using patterns of expression associated with either the nucleic acids or proteins in a computer-based method to diagnose disease.

Diagnostic Methods for Colon Cancer

The present invention also relates to quantitative and qualitative diagnostic assays and methods for detecting, diagnosing, monitoring, staging and predicting cancers by comparing expression of a CSNA or a CSP in a human patient that has or may have colon cancer, or who is at risk of developing colon cancer, with the expression of a CSNA or a CSP in a normal human control. For purposes of the present invention, "expression of a CSNA" or "CSNA expression" means the quantity of CSNA mRNA that can be measured by any method known in the art or the level of transcription that can be measured by any method known in the art in a cell, tissue, organ or whole patient. Similarly, the term "expression of a CSP" or "CSP expression" means the amount of CSP that can be measured by any method known in the art or the level of translation of a CSNA that can be measured by any method known in the art.

The present invention provides methods for diagnosing colon cancer in a patient, in particular adenocarcinoma, by analyzing for changes in levels of CSNA or CSP in cells, tissues, organs or bodily fluids compared with levels of CSNA or CSP in cells, tissues, organs or bodily fluids of preferably the same type from a normal human control, wherein an increase, or decrease in certain cases, in levels of a CSNA or CSP in the patient versus the normal human control is associated with the presence of colon cancer or with a predilection to the disease. In another preferred embodiment, the present invention

provides methods for diagnosing colon cancer in a patient by analyzing changes in the structure of the mRNA of a CSG compared to the mRNA from a normal control. These changes include, without limitation, aberrant splicing, alterations in polyadenylation and/or alterations in 5' nucleotide capping. In yet another preferred embodiment, the present invention provides methods for diagnosing colon cancer in a patient by analyzing changes in a CSP compared to a CSP from a normal patient. These changes include, *e.g.*, alterations, including post translational modifications such as glycosylation and/or phosphorylation of the CSP or changes in the subcellular CSP localization.

For purposes of the present invention, diagnosing means that CSNA or CSP levels are used to determine the presence or absence of disease in a patient. As will be understood by those of skill in the art, measurement of other diagnostic parameters may be required for definitive diagnosis or determination of the appropriate treatment for the disease. The determination may be made by a clinician, a doctor, a testing laboratory, or a patient using an over the counter test. The patient may have symptoms of disease or may be asymptomatic. In addition, the CSNA or CSP levels of the present invention may be used as screening marker to determine whether further tests or biopsies are warranted. In addition, the CSNA or CSP levels may be used to determine the vulnerability or susceptibility to disease.

In a preferred embodiment, the expression of a CSNA is measured by determining the amount of a mRNA that encodes an amino acid sequence selected from SEQ ID NO: 96-237, a homolog, an allelic variant, or a fragment thereof. In a more preferred embodiment, the CSNA expression that is measured is the level of expression of a CSNA mRNA selected from SEQ ID NO: 1-95, or a hybridizing nucleic acid, homologous nucleic acid or allelic variant thereof, or a part of any of these nucleic acid molecules. CSNA expression may be measured by any method known in the art, such as those described *supra*, including measuring mRNA expression by Northern blot, quantitative or qualitative reverse transcriptase PCR (RT-PCR), microarray, dot or slot blots or *in situ* hybridization. *See, e.g.*, Ausubel (1992), *supra*; Ausubel (1999), *supra*; Sambrook (1989), *supra*; and Sambrook (2001), *supra*. CSNA transcription may be measured by any method known in the art including using a reporter gene hooked up to the promoter of a CSG of interest or doing nuclear run-off assays. Alterations in mRNA structure, *e.g.*, aberrant splicing variants, may be determined by any method known in the art, including, RT-PCR followed by sequencing or restriction analysis. As necessary, CSNA expression

may be compared to a known control, such as normal colon nucleic acid, to detect a change in expression.

In another preferred embodiment, the expression of a CSP is measured by determining the level of a CSP having an amino acid sequence selected from the group consisting of SEQ ID NO: 96-237, a homolog, an allelic variant, or a fragment thereof. Such levels are preferably determined in at least one of cells, tissues, organs and/or bodily fluids, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay in accordance with the invention for diagnosing over- or underexpression of a CSNA or CSP compared to normal control bodily fluids, cells, or tissue samples may be used to diagnose the presence of colon cancer. The expression level of a CSP may be determined by any method known in the art, such as those described *supra*. In a preferred embodiment, the CSP expression level may be determined by radioimmunoassays, competitive-binding assays, ELISA, Western blot, FACS, immunohistochemistry, immunoprecipitation, proteomic approaches: two-dimensional gel electrophoresis (2D electrophoresis) and non-gel-based approaches such as mass spectrometry or protein interaction profiling. See, e.g., Harlow (1999), *supra*; Ausubel (1992), *supra*; and Ausubel (1999), *supra*. Alterations in the CSP structure may be determined by any method known in the art, including, e.g., using antibodies that specifically recognize phosphoserine, phosphothreonine or phosphotyrosine residues, two-dimensional polyacrylamide gel electrophoresis (2D PAGE) and/or chemical analysis of amino acid residues of the protein. *Id.*

In a preferred embodiment, a radioimmunoassay (RIA) or an ELISA is used. An antibody specific to a CSP is prepared if one is not already available. In a preferred embodiment, the antibody is a monoclonal antibody. The anti-CSP antibody is bound to a solid support and any free protein binding sites on the solid support are blocked with a protein such as bovine serum albumin. A sample of interest is incubated with the antibody on the solid support under conditions in which the CSP will bind to the anti-CSP antibody. The sample is removed, the solid support is washed to remove unbound material, and an anti-CSP antibody that is linked to a detectable reagent (a radioactive substance for RIA and an enzyme for ELISA) is added to the solid support and incubated under conditions in which binding of the CSP to the labeled antibody will occur. After binding, the unbound labeled antibody is removed by washing. For an ELISA, one or more substrates are added to produce a colored reaction product that is based upon the amount of a CSP in the

sample. For an RIA, the solid support is counted for radioactive decay signals by any method known in the art. Quantitative results for both RIA and ELISA typically are obtained by reference to a standard curve.

Other methods to measure CSP levels are known in the art. For instance, a competition assay may be employed wherein an anti-CSP antibody is attached to a solid support and an allocated amount of a labeled CSP and a sample of interest are incubated with the solid support. The amount of labeled CSP attached to the solid support can be correlated to the quantity of a CSP in the sample.

Of the proteomic approaches, 2D PAGE is a well known technique. Isolation of individual proteins from a sample such as serum is accomplished using sequential separation of proteins by isoelectric point and molecular weight. Typically, polypeptides are first separated by isoelectric point (the first dimension) and then separated by size using an electric current (the second dimension). In general, the second dimension is perpendicular to the first dimension. Because no two proteins with different sequences are identical on the basis of both size and charge, the result of 2D PAGE is a roughly square gel in which each protein occupies a unique spot. Analysis of the spots with chemical or antibody probes, or subsequent protein microsequencing can reveal the relative abundance of a given protein and the identity of the proteins in the sample.

Expression levels of a CSNA can be determined by any method known in the art, including PCR and other nucleic acid methods, such as ligase chain reaction (LCR) and nucleic acid sequence based amplification (NASBA), can be used to detect malignant cells for diagnosis and monitoring of various malignancies. For example, reverse-transcriptase PCR (RT-PCR) is a powerful technique which can be used to detect the presence of a specific mRNA population in a complex mixture of thousands of other mRNA species. In RT-PCR, an mRNA species is first reverse transcribed to complementary DNA (cDNA) with use of the enzyme reverse transcriptase; the cDNA is then amplified as in a standard PCR reaction.

Hybridization to specific DNA molecules (*e.g.*, oligonucleotides) arrayed on a solid support can be used to both detect the expression of and quantitate the level of expression of one or more CSNAs of interest. In this approach, all or a portion of one or more CSNAs is fixed to a substrate. A sample of interest, which may comprise RNA, *e.g.*, total RNA or polyA-selected mRNA, or a complementary DNA (cDNA) copy of the RNA is incubated with the solid support under conditions in which hybridization will occur

between the DNA on the solid support and the nucleic acid molecules in the sample of interest. Hybridization between the substrate-bound DNA and the nucleic acid molecules in the sample can be detected and quantitated by several means, including, without limitation, radioactive labeling or fluorescent labeling of the nucleic acid molecule or a
5 secondary molecule designed to detect the hybrid.

The above tests can be carried out on samples derived from a variety of cells, bodily fluids and/or tissue extracts such as homogenates or solubilized tissue obtained from a patient. Tissue extracts are obtained routinely from tissue biopsy and autopsy material. Bodily fluids useful in the present invention include blood, urine, saliva or any
10 other bodily secretion or derivative thereof. As used herein "blood" includes whole blood, plasma, serum, circulating epithelial cells, constituents, or any derivative of blood.

In addition to detection in bodily fluids, the proteins and nucleic acids of the invention are suitable to detection by cell capture technology. Whole cells may be captured by a variety of methods for example magnetic separation, such as described in U.S.
15 Patent Nos. 5,200,084; 5,186,827; 5,108,933; and 4,925,788, the disclosures of which are incorporated herein by reference in their entireties. Epithelial cells may be captured using such products as Dynabeads® or CELLection™ (DynaL Biotech, Oslo, Norway). Alternatively, fractions of blood may be captured, e.g., the buffy coat fraction (50mm cells isolated from 5ml of blood) containing epithelial cells. In addition, cancer cells may be
20 captured using the techniques described in WO 00/47998, the disclosure of which is incorporated herein by reference in its entirety. Once the cells are captured or concentrated, the proteins or nucleic acids are detected by the means described in the subject application. Alternatively, nucleic acids may be captured directly from blood samples, see U.S. Patent Nos. 6,156,504, 5,501,963; or WO 01/42504, the disclosures of
25 which are incorporated herein by reference in their entireties.

In a preferred embodiment, the specimen tested for expression of CSNA or CSP includes without limitation colon tissue, fecal samples, colonocytes, colon cells grown in cell culture, blood, serum, lymph node tissue, and lymphatic fluid. In another preferred embodiment, especially when metastasis of a primary colon cancer is known or suspected,
30 specimens include, without limitation, tissues from brain, bone, bone marrow, liver, lungs, and adrenal glands. In general, the tissues may be sampled by biopsy, including, without limitation, needle biopsy, e.g., transthoracic needle aspiration, cervical mediastinoscopy,

endoscopic lymph node biopsy, video-assisted thoracoscopy, exploratory thoracotomy, bone marrow biopsy and bone marrow aspiration.

Colonocytes represent an important source of the CSP or CSNA because they provide a picture of the immediate past metabolic history of the GI tract of a subject. In addition, such cells are representative of the cell population from a statistically large sampling frame reflecting the state of the colonic mucosa along the entire length of the colon in a non-invasive manner, in contrast to a limited sampling by colonic biopsy using an invasive procedure involving endoscopy. Specific examples of patents describing the isolation of colonocytes include U.S. Patent Nos. 6,335,193; 6,020,137 5,741,650; 6,258,541; US 2001 0026925 A1; WO 00/63358 A1, the disclosures of which are incorporated herein by reference in their entireties.

All the methods of the present invention may optionally include determining the expression levels of one or more other cancer markers in addition to determining the expression level of a CSNA or CSP. In many cases, the use of another cancer marker will decrease the likelihood of false positives or false negatives. In one embodiment, the one or more other cancer markers include other CSNAs or CSPs as disclosed herein. Other cancer markers useful in the present invention will depend on the cancer being tested and are known to those of skill in the art. In a preferred embodiment, at least one other cancer marker in addition to a particular CSNA or CSP is measured. In a more preferred embodiment, at least two other additional cancer markers are used. In an even more preferred embodiment, at least three, more preferably at least five, even more preferably at least ten additional cancer markers are used.

Diagnosing

In one aspect, the invention provides a method for determining the expression levels and/or structural alterations of one or more CSNA and/or CSP in a sample from a patient suspected of having colon cancer. In general, the method comprises the steps of obtaining the sample from the patient, determining the expression level or structural alterations of a CSNA and/or CSP and then ascertaining whether the patient has colon cancer from the expression level of the CSNA or CSP. In general, if high expression relative to a control of a CSNA or CSP is indicative of colon cancer, a diagnostic assay is considered positive if the level of expression of the CSNA or CSP is at least one and a half times higher, and more preferably are at least two times higher, still more preferably five

times higher, even more preferably at least ten times higher, than in preferably the same cells, tissues or bodily fluid of a normal human control. In contrast, if low expression relative to a control of a CSNA or CSP is indicative of colon cancer, a diagnostic assay is considered positive if the level of expression of the CSNA or CSP is at least one and a half
5 times lower, and more preferably are at least two times lower, still more preferably five times lower, even more preferably at least ten times lower than in preferably the same cells, tissues or bodily fluid of a normal human control. The normal human control may be from a different patient or from uninvolved tissue of the same patient.

The present invention also provides a method of determining whether colon cancer
10 has metastasized in a patient. One may identify whether the colon cancer has metastasized by measuring the expression levels and/or structural alterations of one or more CSNAs and/or CSPs in a variety of tissues. The presence of a CSNA or CSP in a tissue other than colon at levels higher than that of corresponding noncancerous tissue (*e.g.*, the same tissue from another individual) is indicative of metastasis if high level expression of a CSNA or
15 CSP is associated with colon cancer. Similarly, the presence of a CSNA or CSP in a tissue other than colon at levels lower than that of corresponding noncancerous tissue is indicative of metastasis if low level expression of a CSNA or CSP is associated with colon cancer. Further, the presence of a structurally altered CSNA or CSP that is associated with colon cancer is also indicative of metastasis.

20 In general, if high expression relative to a control of a CSNA or CSP is indicative of metastasis, an assay for metastasis is considered positive if the level of expression of the CSNA or CSP is at least one and a half times higher, and more preferably are at least two times higher, still more preferably five times higher, even more preferably at least ten times higher, than in preferably the same cells, tissues or bodily fluid of a normal human
25 control. In contrast, if low expression relative to a control of a CSNA or CSP is indicative of metastasis, an assay for metastasis is considered positive if the level of expression of the CSNA or CSP is at least one and a half times lower, and more preferably are at least two times lower, still more preferably five times lower, even more preferably at least ten times lower than in preferably the same cells, tissues or bodily fluid of a normal human
30 control.

Staging

The invention also provides a method of staging colon cancer in a human patient. The method comprises identifying a human patient having colon cancer and analyzing cells, tissues or bodily fluids from such human patient for expression levels and/or structural alterations of one or more CSNAs or CSPs. First, one or more tumors from a variety of patients are staged according to procedures well known in the art, and the expression levels of one or more CSNAs or CSPs is determined for each stage to obtain a standard expression level for each CSNA and CSP. Then, the CSNA or CSP expression levels of the CSNA or CSP are determined in a biological sample from a patient whose stage of cancer is not known. The CSNA or CSP expression levels from the patient are then compared to the standard expression level. By comparing the expression level of the CSNAs and CSPs from the patient to the standard expression levels, one may determine the stage of the tumor. The same procedure may be followed using structural alterations of a CSNA or CSP to determine the stage of a colon cancer.

Monitoring

Further provided is a method of monitoring colon cancer in a human patient. One may monitor a human patient to determine whether there has been metastasis and, if there has been, when metastasis began to occur. One may also monitor a human patient to determine whether a preneoplastic lesion has become cancerous. One may also monitor a human patient to determine whether a therapy, *e.g.*, chemotherapy, radiotherapy or surgery, has decreased or eliminated the colon cancer. The monitoring may determine if there has been a reoccurrence and, if so, determine its nature. The method comprises identifying a human patient that one wants to monitor for colon cancer, periodically analyzing cells, tissues or bodily fluids from such human patient for expression levels of one or more CSNAs or CSPs, and comparing the CSNA or CSP levels over time to those CSNA or CSP expression levels obtained previously. Patients may also be monitored by measuring one or more structural alterations in a CSNA or CSP that are associated with colon cancer.

If increased expression of a CSNA or CSP is associated with metastasis, treatment failure, or conversion of a preneoplastic lesion to a cancerous lesion, then detecting an increase in the expression level of a CSNA or CSP indicates that the tumor is metastasizing, that treatment has failed or that the lesion is cancerous, respectively. One having ordinary skill in the art would recognize that if this were the case, then a decreased

expression level would be indicative of no metastasis, effective therapy or failure to progress to a neoplastic lesion. If decreased expression of a CSNA or CSP is associated with metastasis, treatment failure, or conversion of a preneoplastic lesion to a cancerous lesion, then detecting a decrease in the expression level of a CSNA or CSP indicates that
5 the tumor is metastasizing, that treatment has failed or that the lesion is cancerous, respectively. In a preferred embodiment, the levels of CSNAs or CSPs are determined from the same cell type, tissue or bodily fluid as prior patient samples. Monitoring a patient for onset of colon cancer metastasis is periodic and preferably is done on a quarterly basis, but may be done more or less frequently.

10 The methods described herein can further be utilized as prognostic assays to identify subjects having or at risk of developing a disease or disorder associated with increased or decreased expression levels of a CSNA and/or CSP. The present invention provides a method in which a test sample is obtained from a human patient and one or more CSNAs and/or CSPs are detected. The presence of higher (or lower) CSNA or CSP
15 levels as compared to normal human controls is diagnostic for the human patient being at risk for developing cancer, particularly colon cancer. The effectiveness of therapeutic agents to decrease (or increase) expression or activity of one or more CSNAs and/or CSPs of the invention can also be monitored by analyzing levels of expression of the CSNAs and/or CSPs in a human patient in clinical trials or in *in vitro* screening assays such as in
20 human cells. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the human patient or cells, as the case may be, to the agent being tested.

Detection of Genetic Lesions or Mutations

The methods of the present invention can also be used to detect genetic lesions or
25 mutations in a CSG, thereby determining if a human with the genetic lesion is susceptible to developing colon cancer or to determine what genetic lesions are responsible, or are partly responsible, for a person's existing colon cancer. Genetic lesions can be detected, for example, by ascertaining the existence of a deletion, insertion and/or substitution of one or more nucleotides from the CSGs of this invention, a chromosomal rearrangement
30 of a CSG, an aberrant modification of a CSG (such as of the methylation pattern of the genomic DNA), or allelic loss of a CSG. Methods to detect such lesions in the CSG of

this invention are known to those having ordinary skill in the art following the teachings of the specification.

Methods of Detecting Noncancerous Colon Diseases

The present invention also provides methods for determining the expression levels
5 and/or structural alterations of one or more CSNAs and/or CSPs in a sample from a patient
suspected of having or known to have a noncancerous colon disease. In general, the
method comprises the steps of obtaining a sample from the patient, determining the
expression level or structural alterations of a CSNA and/or CSP, comparing the expression
level or structural alteration of the CSNA or CSP to a normal colon control, and then
10 ascertaining whether the patient has a noncancerous colon disease. In general, if high
expression relative to a control of a CSNA or CSP is indicative of a particular
noncancerous colon disease, a diagnostic assay is considered positive if the level of
expression of the CSNA or CSP is at least two times higher, and more preferably are at
least five times higher, even more preferably at least ten times higher, than in preferably
15 the same cells, tissues or bodily fluid of a normal human control. In contrast, if low
expression relative to a control of a CSNA or CSP is indicative of a noncancerous colon
disease, a diagnostic assay is considered positive if the level of expression of the CSNA or
CSP is at least two times lower, more preferably are at least five times lower, even more
preferably at least ten times lower than in preferably the same cells, tissues or bodily fluid
20 of a normal human control. The normal human control may be from a different patient or
from uninvolved tissue of the same patient.

One having ordinary skill in the art may determine whether a CSNA and/or CSP is
associated with a particular noncancerous colon disease by obtaining colon tissue from a
patient having a noncancerous colon disease of interest and determining which CSNAs
25 and/or CSPs are expressed in the tissue at either a higher or a lower level than in normal
colon tissue. In another embodiment, one may determine whether a CSNA or CSP
exhibits structural alterations in a particular noncancerous colon disease state by obtaining
colon tissue from a patient having a noncancerous colon disease of interest and
determining the structural alterations in one or more CSNAs and/or CSPs relative to
30 normal colon tissue.

Methods for Identifying Colon Tissue

In another aspect, the invention provides methods for identifying colon tissue. These methods are particularly useful in, *e.g.*, forensic science, colon cell differentiation and development, and in tissue engineering.

5 In one embodiment, the invention provides a method for determining whether a sample is colon tissue or has colon tissue-like characteristics. The method comprises the steps of providing a sample suspected of comprising colon tissue or having colon tissue-like characteristics, determining whether the sample expresses one or more CSNAs and/or CSPs, and, if the sample expresses one or more CSNAs and/or CSPs, concluding that the
10 sample comprises colon tissue. In a preferred embodiment, the CSNA encodes a polypeptide having an amino acid sequence selected from SEQ ID NO: 96-237, or a homolog, allelic variant or fragment thereof. In a more preferred embodiment, the CSNA has a nucleotide sequence selected from SEQ ID NO: 1-95, or a hybridizing nucleic acid, an allelic variant or a part thereof. Determining whether a sample expresses a CSNA can
15 be accomplished by any method known in the art. Preferred methods include hybridization to microarrays, Northern blot hybridization, and quantitative or qualitative RT-PCR. In another preferred embodiment, the method can be practiced by determining whether a CSP is expressed. Determining whether a sample expresses a CSP can be accomplished by any method known in the art. Preferred methods include Western blot,
20 ELISA, RIA and 2D PAGE. In one embodiment, the CSP has an amino acid sequence selected from SEQ ID NO: 96-237, or a homolog, allelic variant or fragment thereof. In another preferred embodiment, the expression of at least two CSNAs and/or CSPs is determined. In a more preferred embodiment, the expression of at least three, more preferably four and even more preferably five CSNAs and/or CSPs are determined.

25 In one embodiment, the method can be used to determine whether an unknown tissue is colon tissue. This is particularly useful in forensic science, in which small, damaged pieces of tissues that are not identifiable by microscopic or other means are recovered from a crime or accident scene. In another embodiment, the method can be used to determine whether a tissue is differentiating or developing into colon tissue. This
30 is important in monitoring the effects of the addition of various agents to cell or tissue culture, *e.g.*, in producing new colon tissue by tissue engineering. These agents include, *e.g.*, growth and differentiation factors, extracellular matrix proteins and culture medium. Other factors that may be measured for effects on tissue development and differentiation

include gene transfer into the cells or tissues, alterations in pH, aqueous:air interface and various other culture conditions.

Methods for Producing and Modifying Colon Tissue

In another aspect, the invention provides methods for producing engineered colon
5 tissue or cells. In one embodiment, the method comprises the steps of providing cells, introducing a CSNA or a CSG into the cells, and growing the cells under conditions in which they exhibit one or more properties of colon tissue cells. In a preferred embodiment, the cells are pluripotent. As is well known in the art, normal colon tissue comprises a large number of different cell types. Thus, in one embodiment, the
10 engineered colon tissue or cells comprises one of these cell types. In another embodiment, the engineered colon tissue or cells comprises more than one colon cell type. Further, the culture conditions of the cells or tissue may require manipulation in order to achieve full differentiation and development of the colon cell tissue. Methods for manipulating culture conditions are well known in the art.

15 Nucleic acid molecules encoding one or more CSPs are introduced into cells, preferably pluripotent cells. In a preferred embodiment, the nucleic acid molecules encode CSPs having amino acid sequences selected from SEQ ID NO: 96-237, or homologous proteins, analogs, allelic variants or fragments thereof. In a more preferred embodiment, the nucleic acid molecules have a nucleotide sequence selected from SEQ ID
20 NO: 1-95, or hybridizing nucleic acids, allelic variants or parts thereof. In another highly preferred embodiment, a CSG is introduced into the cells. Expression vectors and methods of introducing nucleic acid molecules into cells are well known in the art and are described in detail, *supra*.

Artificial colon tissue may be used to treat patients who have lost some or all of
25 their colon function.

Pharmaceutical Compositions

In another aspect, the invention provides pharmaceutical compositions comprising the nucleic acid molecules, polypeptides, fusion proteins, antibodies, antibody derivatives, antibody fragments, agonists, antagonists, or inhibitors of the present invention. In a
30 preferred embodiment, the pharmaceutical composition comprises a CSNA or part thereof. In a more preferred embodiment, the CSNA has a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-95, a nucleic acid that hybridizes thereto, an allelic

variant thereof, or a nucleic acid that has substantial sequence identity thereto. In another preferred embodiment, the pharmaceutical composition comprises a CSP or fragment thereof. In a more preferred embodiment, the pharmaceutical composition comprises a CSP having an amino acid sequence that is selected from the group consisting of SEQ ID NO: 96-237, a polypeptide that is homologous thereto, a fusion protein comprising all or a portion of the polypeptide, or an analog or derivative thereof. In another preferred embodiment, the pharmaceutical composition comprises an anti-CSP antibody, preferably an antibody that specifically binds to a CSP having an amino acid that is selected from the group consisting of SEQ ID NO: 96-237, or an antibody that binds to a polypeptide that is homologous thereto, a fusion protein comprising all or a portion of the polypeptide, or an analog or derivative thereof.

Due to the association of angiogenesis with cancer vascularization there is great need of new markers and methods for diagnosing angiogenesis activity to identify developing tumors and angiogenesis related diseases. Furthermore, great need is also present for new molecular targets useful in the treatment of angiogenesis and angiogenesis related diseases such as cancer. In addition known modulators of angiogenesis such as endostatin or vascular endothelial growth factor (VEGF). Use of the methods and compositions disclosed herein in combination with anti-angiogenesis drugs, drugs that block the matrix breakdown (such as BMS-275291, Dalteparin (Fragmin®), Suramin), drugs that inhibit endothelial cells (2-methoxyestradiol (2-ME), CC-5013 (Thalidomide Analog), Combretastatin A4 Phosphate, LY317615 (Protein Kinase C Beta Inhibitor), Soy Isoflavone (Genistein; Soy Protein Isolate), Thalidomide), drugs that block activators of angiogenesis (AE-941 (Neovastat™; GW786034), Anti-VEGF Antibody (Bevacizumab; Avastin™), Interferon-alpha, PTK787/ZK 222584, VEGF-Trap, ZD6474), Drugs that inhibit endothelial-specific integrin/survival signaling (EMD 121974, Anti-Anb3 Integrin Antibody (Medi-522; Vitaxin™)).

Such a composition typically contains from about 0.1 to 90% by weight of a therapeutic agent of the invention formulated in and/or with a pharmaceutically acceptable carrier or excipient.

Pharmaceutical formulation is a well-established art that is further described in Gennaro (ed.), Remington: The Science and Practice of Pharmacy, 20th ed., Lippincott, Williams & Wilkins (2000); Ansel *et al.*, Pharmaceutical Dosage Forms and Drug Delivery Systems, 7th ed., Lippincott Williams & Wilkins (1999); and Kibbe (ed.),

Handbook of Pharmaceutical Excipients American Pharmaceutical Association, 3rd ed. (2000) and thus need not be described in detail herein.

Briefly, formulation of the pharmaceutical compositions of the present invention will depend upon the route chosen for administration. The pharmaceutical compositions
5 utilized in this invention can be administered by various routes including both enteral and parenteral routes, including oral, intravenous, intramuscular, subcutaneous, inhalation, topical, sublingual, rectal, intra-arterial, intramedullary, intrathecal, intraventricular, transmucosal, transdermal, intranasal, intraperitoneal, intrapulmonary, and intrauterine.

Oral dosage forms can be formulated as tablets, pills, dragees, capsules, liquids,
10 gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Solid formulations of the compositions for oral administration can contain suitable carriers or excipients, such as carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, sodium
15 carboxymethylcellulose, or microcrystalline cellulose; gums including arabic and tragacanth; proteins such as gelatin and collagen; inorganics, such as kaolin, calcium carbonate, dicalcium phosphate, sodium chloride; and other agents such as acacia and alginic acid.

Agents that facilitate disintegration and/or solubilization can be added, such as the
20 cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate, microcrystalline cellulose, cornstarch, sodium starch glycolate, and alginic acid.

Tablet binders that can be used include acacia, methylcellulose, sodium carboxymethylcellulose, polyvinylpyrrolidone (Povidone™), hydroxypropyl methylcellulose, sucrose, starch and ethylcellulose.

25 Lubricants that can be used include magnesium stearates, stearic acid, silicone fluid, talc, waxes, oils, and colloidal silica.

Fillers, agents that facilitate disintegration and/or solubilization, tablet binders and lubricants, including the aforementioned, can be used singly or in combination.

Solid oral dosage forms need not be uniform throughout. For example, dragee
30 cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which can also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures.

Oral dosage forms of the present invention include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Additionally, dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, *i.e.*, dosage.

Liquid formulations of the pharmaceutical compositions for oral (enteral) administration are prepared in water or other aqueous vehicles and can contain various suspending agents such as methylcellulose, alginates, tragacanth, pectin, kelgin, carrageenan, acacia, polyvinylpyrrolidone, and polyvinyl alcohol. The liquid formulations can also include solutions, emulsions, syrups and elixirs containing, together with the active compound(s), wetting agents, sweeteners, and coloring and flavoring agents.

The pharmaceutical compositions of the present invention can also be formulated for parenteral administration. Formulations for parenteral administration can be in the form of aqueous or non-aqueous isotonic sterile injection solutions or suspensions.

For intravenous injection, water soluble versions of the compounds of the present invention are formulated in, or if provided as a lyophilate, mixed with, a physiologically acceptable fluid vehicle, such as 5% dextrose ("D5"), physiologically buffered saline, 0.9% saline, Hanks' solution, or Ringer's solution. Intravenous formulations may include carriers, excipients or stabilizers including, without limitation, calcium, human serum albumin, citrate, acetate, calcium chloride, carbonate, and other salts.

Intramuscular preparations, *e.g.* a sterile formulation of a suitable soluble salt form of the compounds of the present invention, can be dissolved and administered in a pharmaceutical excipient such as Water-for-Injection, 0.9% saline, or 5% glucose solution. Alternatively, a suitable insoluble form of the compound can be prepared and administered as a suspension in an aqueous base or a pharmaceutically acceptable oil base, such as an ester of a long chain fatty acid (*e.g.*, ethyl oleate), fatty oils such as sesame oil, triglycerides, or liposomes.

Parenteral formulations of the compositions can contain various carriers such as vegetable oils, dimethylacetamide, dimethylformamide, ethyl lactate, ethyl carbonate,

isopropyl myristate, ethanol, polyols (glycerol, propylene glycol, liquid polyethylene glycol, and the like).

Aqueous injection suspensions can also contain substances that increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran.

- 5 Non-lipid polycationic amino polymers can also be used for delivery. Optionally, the suspension can also contain suitable stabilizers or agents that increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

- Pharmaceutical compositions of the present invention can also be formulated to permit injectable, long-term, deposition. Injectable depot forms may be made by forming
10 microencapsulated matrices of the compound in biodegradable polymers such as polylactide-polyglycolide. Depending upon the ratio of drug to polymer and the nature of the particular polymer employed, the rate of drug release can be controlled. Examples of other biodegradable polymers include poly(orthoesters) and poly(anhydrides). Depot injectable formulations are also prepared by entrapping the drug in microemulsions that
15 are compatible with body tissues.

- The pharmaceutical compositions of the present invention can be administered topically. For topical use the compounds of the present invention can also be prepared in suitable forms to be applied to the skin, or mucus membranes of the nose and throat, and can take the form of lotions, creams, ointments, liquid sprays or inhalants, drops, tinctures,
20 lozenges, or throat paints. Such topical formulations further can include chemical compounds such as dimethylsulfoxide (DMSO) to facilitate surface penetration of the active ingredient. In other transdermal formulations, typically in patch-delivered formulations, the pharmaceutically active compound is formulated with one or more skin penetrants, such as 2-N-methyl-pyrrolidone (NMP) or Azone. A topical semi-solid
25 ointment formulation typically contains a concentration of the active ingredient from about 1 to 20%, e.g., 5 to 10%, in a carrier such as a pharmaceutical cream base.

For application to the eyes or ears, the compounds of the present invention can be presented in liquid or semi-liquid form formulated in hydrophobic or hydrophilic bases as ointments, creams, lotions, paints or powders.

- 30 For rectal administration the compounds of the present invention can be administered in the form of suppositories admixed with conventional carriers such as cocoa butter, wax or other glyceride.

Inhalation formulations can also readily be formulated. For inhalation, various powder and liquid formulations can be prepared. For aerosol preparations, a sterile formulation of the compound or salt form of the compound may be used in inhalers, such as metered dose inhalers, and nebulizers. Aerosolized forms may be especially useful for treating respiratory disorders.

Alternatively, the compounds of the present invention can be in powder form for reconstitution in the appropriate pharmaceutically acceptable carrier at the time of delivery.

The pharmaceutically active compound in the pharmaceutical compositions of the present invention can be provided as the salt of a variety of acids, including but not limited to hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms.

After pharmaceutical compositions have been prepared, they are packaged in an appropriate container and labeled for treatment of an indicated condition.

The active compound will be present in an amount effective to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

A "therapeutically effective dose" refers to that amount of active ingredient, for example CSP polypeptide, fusion protein, or fragments thereof, antibodies specific for CSP, agonists, antagonists or inhibitors of CSP, which ameliorates the signs or symptoms of the disease or prevent progression thereof; as would be understood in the medical arts, cure, although desired, is not required.

The therapeutically effective dose of the pharmaceutical agents of the present invention can be estimated initially by *in vitro* tests, such as cell culture assays, followed by assay in model animals, usually mice, rats, rabbits, dogs, or pigs. The animal model can also be used to determine an initial preferred concentration range and route of administration.

For example, the ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population) can be determined in one or more cell culture of animal model systems. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as LD50/ED50. Pharmaceutical compositions that exhibit large therapeutic indices are preferred.

The data obtained from cell culture assays and animal studies are used in formulating an initial dosage range for human use, and preferably provide a range of circulating concentrations that includes the ED50 with little or no toxicity. After administration, or between successive administrations, the circulating concentration of active agent varies within this range depending upon pharmacokinetic factors well known in the art, such as the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors specific to the subject requiring treatment. Factors that can be taken into account by the practitioner include the severity of the disease state, general health of the subject, age, weight, gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Where the therapeutic agent is a protein or antibody of the present invention, the therapeutic protein or antibody agent typically is administered at a daily dosage of 0.01 mg to 30 mg/kg of body weight of the patient (*e.g.*, 1 mg/kg to 5 mg/kg). The pharmaceutical formulation can be administered in multiple doses per day, if desired, to achieve the total desired daily dose.

Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

Conventional methods, known to those of ordinary skill in the art of medicine, can be used to administer the pharmaceutical formulation(s) of the present invention to the patient. The pharmaceutical compositions of the present invention can be administered alone, or in combination with other therapeutic agents or interventions.

30 Therapeutic Methods

The present invention further provides methods of treating subjects having defects in a gene of the invention, *e.g.*, in expression, activity, distribution, localization, and/or

solubility, which can manifest as a disorder of colon function. As used herein, "treating" includes all medically-acceptable types of therapeutic intervention, including palliation and prophylaxis (prevention) of disease. The term "treating" encompasses any improvement of a disease, including minor improvements. These methods are discussed
5 below.

Gene Therapy and Vaccines

The isolated nucleic acids of the present invention can also be used to drive *in vivo* expression of the polypeptides of the present invention. *In vivo* expression can be driven from a vector, typically a viral vector, often a vector based upon a replication incompetent
10 retrovirus, an adenovirus, or an adeno-associated virus (AAV), for the purpose of gene therapy. *In vivo* expression can also be driven from signals endogenous to the nucleic acid or from a vector, often a plasmid vector, such as pVAX1 (Invitrogen, Carlsbad, CA, USA), for purpose of "naked" nucleic acid vaccination, as further described in U.S. Patent Nos. 5,589,466; 5,679,647; 5,804,566; 5,830,877; 5,843,913; 5,880,104; 5,958,891;
15 5,985,847; 6,017,897; 6,110,898; 6,204,250, the disclosures of which are incorporated herein by reference in their entireties. For cancer therapy, it is preferred that the vector also be tumor-selective. See, e.g., Doronin *et al.*, *J. Virol.* 75: 3314-24 (2001).

In another embodiment of the therapeutic methods of the present invention, a therapeutically effective amount of a pharmaceutical composition comprising a nucleic
20 acid molecule of the present invention is administered. The nucleic acid molecule can be delivered in a vector that drives expression of a CSP, fusion protein, or fragment thereof, or without such vector. Nucleic acid compositions that can drive expression of a CSP are administered, for example, to complement a deficiency in the native CSP, or as DNA vaccines. Expression vectors derived from virus, replication deficient retroviruses,
25 adenovirus, adeno-associated (AAV) virus, herpes virus, or vaccinia virus can be used as can plasmids. See, e.g., Cid-Arregui, *supra*. In a preferred embodiment, the nucleic acid molecule encodes a CSP having the amino acid sequence of SEQ ID NO: 96-237, or a fragment, fusion protein, allelic variant or homolog thereof.

In still other therapeutic methods of the present invention, pharmaceutical
30 compositions comprising host cells that express a CSP, fusions, or fragments thereof can be administered. In such cases, the cells are typically autologous, so as to circumvent xenogeneic or allotypic rejection, and are administered to complement defects in CSP

production or activity. In a preferred embodiment, the nucleic acid molecules in the cells encode a CSP having the amino acid sequence of SEQ ID NO: 96-237, or a fragment, fusion protein, allelic variant or homolog thereof.

Antisense Administration

5 Antisense nucleic acid compositions, or vectors that drive expression of a CSG antisense nucleic acid, are administered to downregulate transcription and/or translation of a CSG in circumstances in which excessive production, or production of aberrant protein, is the pathophysiologic basis of disease.

Antisense compositions useful in therapy can have a sequence that is
10 complementary to coding or to noncoding regions of a CSG. For example, oligonucleotides derived from the transcription initiation site, *e.g.*, between positions -10 and +10 from the start site, are preferred.

Catalytic antisense compositions, such as ribozymes, that are capable of sequence-specific hybridization to CSG transcripts, are also useful in therapy. *See, e.g.*,
15 Phylactou, *Adv. Drug Deliv. Rev.* 44(2-3): 97-108 (2000); Phylactou *et al.*, *Hum. Mol. Genet.* 7(10): 1649-53 (1998); Rossi, *Ciba Found. Symp.* 209: 195-204 (1997); and Sigurdsson *et al.*, *Trends Biotechnol.* 13(8): 286-9 (1995).

Other nucleic acids useful in the therapeutic methods of the present invention are those that are capable of triplex helix formation in or near the CSG genomic locus. Such
20 triplexing oligonucleotides are able to inhibit transcription. *See, e.g.*, Intody *et al.*, *Nucleic Acids Res.* 28(21): 4283-90 (2000); and McGuffie *et al.*, *Cancer Res.* 60(14): 3790-9 (2000). Pharmaceutical compositions comprising such triplex forming oligos (TFOs) are administered in circumstances in which excessive production, or production of aberrant protein, is a pathophysiologic basis of disease.

25 In a preferred embodiment, the antisense molecule is derived from a nucleic acid molecule encoding a CSP, preferably a CSP comprising an amino acid sequence of SEQ ID NO: 96-237, or a fragment, allelic variant or homolog thereof. In a more preferred embodiment, the antisense molecule is derived from a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-95, or a part, allelic variant, substantially similar or
30 hybridizing nucleic acid thereof.

Polypeptide Administration

In one embodiment of the therapeutic methods of the present invention, a therapeutically effective amount of a pharmaceutical composition comprising a CSP, a fusion protein, fragment, analog or derivative thereof is administered to a subject with a clinically-significant CSP defect.

5 Protein compositions are administered, for example, to complement a deficiency in native CSP. In other embodiments, protein compositions are administered as a vaccine to elicit a humoral and/or cellular immune response to CSP. The immune response can be used to modulate activity of CSP or, depending on the immunogen, to immunize against aberrant or aberrantly expressed forms, such as mutant or inappropriately expressed
10 isoforms. In yet other embodiments, protein fusions having a toxic moiety are administered to ablate cells that aberrantly accumulate CSP.

In a preferred embodiment, the polypeptide administered is a CSP comprising an amino acid sequence of SEQ ID NO: 96-237, or a fusion protein, allelic variant, homolog, analog or derivative thereof. In a more preferred embodiment, the polypeptide is encoded
15 by a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-95, or a part, allelic variant, substantially similar or hybridizing nucleic acid thereof.

Antibody, Agonist and Antagonist Administration

In another embodiment of the therapeutic methods of the present invention, a therapeutically effective amount of a pharmaceutical composition comprising an antibody
20 (including fragment or derivative thereof) of the present invention is administered. As is well known, antibody compositions are administered, for example, to antagonize activity of CSP, or to target therapeutic agents to sites of CSP presence and/or accumulation. In a preferred embodiment, the antibody specifically binds to a CSP comprising an amino acid sequence of SEQ ID NO: 96-237, or a fusion protein, allelic variant, homolog, analog or
25 derivative thereof. In a more preferred embodiment, the antibody specifically binds to a CSP encoded by a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-95, or a part, allelic variant, substantially similar or hybridizing nucleic acid thereof.

The present invention also provides methods for identifying modulators which bind to a CSP or have a modulatory effect on the expression or activity of a CSP.
30 Modulators which decrease the expression or activity of CSP (antagonists) are believed to be useful in treating colon cancer. Such screening assays are known to those of skill in the art and include, without limitation, cell-based assays and cell-free assays. Small molecules

predicted via computer imaging to specifically bind to regions of a CSP can also be designed, synthesized and tested for use in the imaging and treatment of colon cancer. Further, libraries of molecules can be screened for potential anticancer agents by assessing the ability of the molecule to bind to the CSPs identified herein. Molecules identified in the library as being capable of binding to a CSP are key candidates for further evaluation for use in the treatment of colon cancer. In a preferred embodiment, these molecules will downregulate expression and/or activity of a CSP in cells.

In another embodiment of the therapeutic methods of the present invention, a pharmaceutical composition comprising a non-antibody antagonist of CSP is administered. Antagonists of CSP can be produced using methods generally known in the art. In particular, purified CSP can be used to screen libraries of pharmaceutical agents, often combinatorial libraries of small molecules, to identify those that specifically bind and antagonize at least one activity of a CSP.

In other embodiments a pharmaceutical composition comprising an agonist of a CSP is administered. Agonists can be identified using methods analogous to those used to identify antagonists.

In a preferred embodiment, the antagonist or agonist specifically binds to and antagonizes or agonizes, respectively, a CSP comprising an amino acid sequence of SEQ ID NO: 96-237, or a fusion protein, allelic variant, homolog, analog or derivative thereof. In a more preferred embodiment, the antagonist or agonist specifically binds to and antagonizes or agonizes, respectively, a CSP encoded by a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-95, or a part, allelic variant, substantially similar or hybridizing nucleic acid thereof.

25 *Targeting Colon Tissue*

The invention also provides a method in which a polypeptide of the invention, or an antibody thereto, is linked to a therapeutic agent such that it can be delivered to the colon or to specific cells in the colon. In a preferred embodiment, an anti-CSP antibody is linked to a therapeutic agent and is administered to a patient in need of such therapeutic agent. The therapeutic agent may be a toxin, if colon tissue needs to be selectively destroyed. This would be useful for targeting and killing colon cancer cells. In another embodiment, the therapeutic agent may be a growth or differentiation factor, which would be useful for promoting colon cell function.

In another embodiment, an anti-CSP antibody may be linked to an imaging agent that can be detected using, *e.g.*, magnetic resonance imaging, CT or PET. This would be useful for determining and monitoring colon function, identifying colon cancer tumors, and identifying noncancerous colon diseases.

5

EXAMPLES

Example 1a: Alternative Splice Variants

We identified gene transcripts using the Gencarta™ tools (Compugen Ltd., Tel Aviv, Israel) and a variety of public and proprietary databases. These splice variants are either sequences which differ from a previously defined sequence or new uses of known sequences. In general related variants are annotated as DEX0448_XXX.nt.1, DEX0448_XXX.nt.2, DEX0448_XXX.nt.3, etc. The variant DNA sequences encode proteins which differ from a previously defined protein sequence. In relation to the nucleotide sequence naming convention, protein variants are annotated as DEX0448_XXX.aa.1, DEX0448_XXX.aa.2, etc., wherein transcript DEX0448_XXX.nt.1 encodes protein DEX0448_XXX.aa.1. A single transcript may encode a protein from an alternate Open Reading Frame (ORF) which is designated DEX0448_XXX.orf.1. Additionally, multiple transcripts may encode for a single protein. In this case, DEX0448_XXX.nt.1 and DEX0448_XXX.nt.2 will both be associated with DEX0448_XXX.aa.1.

The mapping of the nucleic acid ("NT") SEQ ID NO; DEX ID; chromosomal location (if known); open reading frame (ORF) location; amino acid ("AA") SEQ ID NO; AA DEX ID; are shown in the table below.

SEQ ID NO	DEX ID	Chromo Map	ORF Loc	SEQ ID NO	DEX ID
1	DEX0448_001.nt.1	6p12.2	1-547	96	DEX0448_001.aa.1
1	DEX0448_001.nt.1	6p12.2	212-538	97	DEX0448_001.orf.1
2	DEX0448_002.nt.1	X;22899334-22902697	1-323	98	DEX0448_002.aa.1
2	DEX0448_002.nt.1	X;22899334-22902697	86-310	99	DEX0448_002.orf.1
3	DEX0448_003.nt.1	8p23.1	250-924	100	DEX0448_003.aa.1
3	DEX0448_003.nt.1	8p23.1	3-545	101	DEX0448_003.orf.1
4	DEX0448_004.nt.1	19q13.2	1-686	102	DEX0448_004.aa.1
4	DEX0448_004.nt.1	19q13.2	1-666	103	DEX0448_004.orf.1
5	DEX0448_005.nt.1	7q32.3	206-431	104	DEX0448_005.aa.1

5	DEX0448_005.nt.1	7q32.3	173-427	105	DEX0448_005.orf.1
6	DEX0448_005.nt.2	7q32.3	206-431	104	DEX0448_005.aa.1
6	DEX0448_005.nt.2	7q32.3	173-427	106	DEX0448_005.orf.2
7	DEX0448_006.nt.1	1q32.2	1-200	107	DEX0448_006.aa.1
7	DEX0448_006.nt.1	1q32.2	3-200	108	DEX0448_006.orf.1
8	DEX0448_007.nt.1	1q22	61-443	109	DEX0448_007.aa.1
8	DEX0448_007.nt.1	1q22	3-353	110	DEX0448_007.orf.1
9	DEX0448_008.nt.1	21q22.2	172-676	111	DEX0448_008.aa.1
9	DEX0448_008.nt.1	21q22.2	2-676	112	DEX0448_008.orf.1
10	DEX0448_009.nt.1	17q23.3	3-527	113	DEX0448_009.aa.1
11	DEX0448_010.nt.1	5q32	60-869	114	DEX0448_010.aa.1
12	DEX0448_011.nt.1	17q21.32	1-673	115	DEX0448_011.aa.1
12	DEX0448_011.nt.1	17q21.32	311-673	116	DEX0448_011.orf.1
13	DEX0448_012.nt.1	1p13.3	6-207	117	DEX0448_012.aa.1
13	DEX0448_012.nt.1	1p13.3	18-260	118	DEX0448_012.orf.1
14	DEX0448_013.nt.1	11q24.2	157-911	119	DEX0448_013.aa.1
14	DEX0448_013.nt.1	11q24.2	279-887	120	DEX0448_013.orf.1
15	DEX0448_014.nt.1	11q13.4	49-431	121	DEX0448_014.aa.1
15	DEX0448_014.nt.1	11q13.4	3-431	122	DEX0448_014.orf.1
16	DEX0448_015.nt.1	11q13.4	1-230	123	DEX0448_015.aa.1
16	DEX0448_015.nt.1	11q13.4	232-561	124	DEX0448_015.orf.1
17	DEX0448_016.nt.1	15q25.3	4-595	125	DEX0448_016.aa.1
18	DEX0448_016.nt.2	15q25.3	209-833	126	DEX0448_016.aa.2
18	DEX0448_016.nt.2	15q25.3	290-829	127	DEX0448_016.orf.2
19	DEX0448_016.nt.3	15q25.3	54-503	128	DEX0448_016.aa.3
20	DEX0448_017.nt.1	8q24.3	1-898	129	DEX0448_017.aa.1
20	DEX0448_017.nt.1	8q24.3	80-892	130	DEX0448_017.orf.1
21	DEX0448_018.nt.1	6p21.1	133-1745	131	DEX0448_018.aa.1
21	DEX0448_018.nt.1	6p21.1	842-1411	132	DEX0448_018.orf.1
22	DEX0448_019.nt.1	16q23.1	46-378	133	DEX0448_019.aa.1
23	DEX0448_020.nt.1	3q27.2	1-783	134	DEX0448_020.aa.1
24	DEX0448_021.nt.1	11p15.5	353-1435	135	DEX0448_021.aa.1
25	DEX0448_021.nt.2	11p15.5	353-1339	136	DEX0448_021.aa.2
26	DEX0448_021.nt.3	11p15.5	1-492	137	DEX0448_021.aa.3
27	DEX0448_022.nt.1	6_DR51;3849138-3852035	2-203	138	DEX0448_022.aa.1
27	DEX0448_022.nt.1	6_DR51;3849138-3852035	214-618	139	DEX0448_022.orf.1
28	DEX0448_022.nt.2	6_DR51;3849138-3851915	149-497	140	DEX0448_022.aa.2
28	DEX0448_022.nt.2	6_DR51;3849138-3851915	109-513	141	DEX0448_022.orf.2
29	DEX0448_023.nt.1	19q13.2	115-777	142	DEX0448_023.aa.1
29	DEX0448_023.nt.1	19q13.2	22-882	143	DEX0448_023.orf.1
30	DEX0448_023.nt.2	19q13.2	387-831	144	DEX0448_023.aa.2
30	DEX0448_023.nt.2	19q13.2	1337-1786	145	DEX0448_023.orf.2
31	DEX0448_023.nt.3	10q26.3	750-3186	146	DEX0448_023.aa.3
31	DEX0448_023.nt.3	10q26.3	742-2067	147	DEX0448_023.orf.3

32	DEX0448_024.nt.1	19p13.12	37-1091	148	DEX0448_024.aa.1
32	DEX0448_024.nt.1	19p13.12	1-669	149	DEX0448_024.orf.1
33	DEX0448_025.nt.1	11p15.5	1-783	150	DEX0448_025.aa.1
33	DEX0448_025.nt.1	11p15.5	3-779	151	DEX0448_025.orf.1
34	DEX0448_026.nt.1	19q13.2	179-2128	152	DEX0448_026.aa.1
35	DEX0448_026.nt.2	19q13.2	341-1504	153	DEX0448_026.aa.2
36	DEX0448_027.nt.1	3p26.1	2590-2842	154	DEX0448_027.aa.1
36	DEX0448_027.nt.1	3p26.1	3377-4513	155	DEX0448_027.orf.1
37	DEX0448_027.nt.2	3p26.1	2590-2842	154	DEX0448_027.aa.1
37	DEX0448_027.nt.2	3p26.1	3245-4381	156	DEX0448_027.orf.2
38	DEX0448_027.nt.3	3p26.1	2590-3667	157	DEX0448_027.aa.3
38	DEX0448_027.nt.3	3p26.1	2677-3663	158	DEX0448_027.orf.3
39	DEX0448_027.nt.4	3p26.1	2590-2842	154	DEX0448_027.aa.1
39	DEX0448_027.nt.4	3p26.1	2970-4244	159	DEX0448_027.orf.4
40	DEX0448_027.nt.5	3p26.1	2590-2842	154	DEX0448_027.aa.1
40	DEX0448_027.nt.5	3p26.1	2970-4244	160	DEX0448_027.orf.5
41	DEX0448_027.nt.6	3p26.1	1-195	161	DEX0448_027.aa.6
41	DEX0448_027.nt.6	3p26.1	354-686	162	DEX0448_027.orf.6
42	DEX0448_028.nt.1	6p21.32	436-874	163	DEX0448_028.aa.1
42	DEX0448_028.nt.1	6p21.32	61-870	164	DEX0448_028.orf.1
43	DEX0448_028.nt.2	6p21.32	186-624	163	DEX0448_028.aa.1
43	DEX0448_028.nt.2	6p21.32	80-619	165	DEX0448_028.orf.2
44	DEX0448_029.nt.1	19q13.33	101-2488	166	DEX0448_029.aa.1
45	DEX0448_029.nt.2	19q13.33	163-2043	167	DEX0448_029.aa.2
46	DEX0448_029.nt.3	19q13.33	166-2046	168	DEX0448_029.aa.3
47	DEX0448_029.nt.4	19q13.33	166-1614	169	DEX0448_029.aa.4
48	DEX0448_029.nt.5	19q13.33	38-1852	170	DEX0448_029.aa.5
49	DEX0448_029.nt.6	19q13.33	38-1522	171	DEX0448_029.aa.6
50	DEX0448_029.nt.7	19q13.33	26-1633	172	DEX0448_029.orf.7
50	DEX0448_029.nt.7	19q13.33	1-1635	173	DEX0448_029.aa.7
51	DEX0448_030.nt.1	1q25.3	2-1447	174	DEX0448_030.orf.1
51	DEX0448_030.nt.1	1q25.3	91-1455	175	DEX0448_030.aa.1
52	DEX0448_031.nt.1	1q24.2	324-1343	176	DEX0448_031.aa.1
53	DEX0448_032.nt.1	8q24.3	46-957	177	DEX0448_032.orf.1
53	DEX0448_032.nt.1	8q24.3	380-938	178	DEX0448_032.aa.1
54	DEX0448_033.nt.1	18p11.32	3-1454	179	DEX0448_033.orf.1
54	DEX0448_033.nt.1	18p11.32	3-1455	180	DEX0448_033.aa.1
55	DEX0448_033.nt.2	18p11.32	3-1454	181	DEX0448_033.orf.2
55	DEX0448_033.nt.2	18p11.32	3-1455	180	DEX0448_033.aa.1
56	DEX0448_033.nt.3	18p11.32	3-1454	182	DEX0448_033.orf.3
56	DEX0448_033.nt.3	18p11.32	3-1455	180	DEX0448_033.aa.1
57	DEX0448_033.nt.4	18p11.32	1199-1945	183	DEX0448_033.orf.4
57	DEX0448_033.nt.4	18p11.32	3-669	184	DEX0448_033.aa.4
58	DEX0448_033.nt.5	18p11.32	3-1250	185	DEX0448_033.orf.5
58	DEX0448_033.nt.5	18p11.32	3-1251	186	DEX0448_033.aa.5
59	DEX0448_033.nt.6	18p11.32	3-1454	187	DEX0448_033.orf.6
59	DEX0448_033.nt.6	18p11.32	3-1455	180	DEX0448_033.aa.1

60	DEX0448_033.nt.7	18p11.32	3-1049	188	DEX0448_033.orf.7
60	DEX0448_033.nt.7	18p11.32	134-1052	189	DEX0448_033.aa.7
61	DEX0448_033.nt.8	18p11.32	3-1454	190	DEX0448_033.orf.8
61	DEX0448_033.nt.8	18p11.32	3-1455	180	DEX0448_033.aa.1
62	DEX0448_033.nt.9	18p11.32	3-1454	191	DEX0448_033.orf.9
62	DEX0448_033.nt.9	18p11.32	3-1455	180	DEX0448_033.aa.1
63	DEX0448_033.nt.10	18p11.32	3-1454	192	DEX0448_033.orf.10
63	DEX0448_033.nt.10	18p11.32	3-1455	180	DEX0448_033.aa.1
64	DEX0448_034.nt.1	15q25.3	1-417	193	DEX0448_034.aa.1
65	DEX0448_035.nt.1	1q42.3	61-1218	194	DEX0448_035.aa.1
66	DEX0448_035.nt.2	1q42.3	61-1221	194	DEX0448_035.aa.1
67	DEX0448_035.nt.3	1q42.3	61-1221	194	DEX0448_035.aa.1
68	DEX0448_035.nt.4	1q42.3	61-1536	195	DEX0448_035.aa.4
69	DEX0448_036.nt.1	4p16.3	407-1480	196	DEX0448_036.aa.1
70	DEX0448_036.nt.2	4p16.3	167-1262	197	DEX0448_036.aa.2
70	DEX0448_036.nt.2	4p16.3	466-1041	198	DEX0448_036.orf.2
71	DEX0448_036.nt.3	4p16.3	31-569	199	DEX0448_036.aa.3
71	DEX0448_036.nt.3	4p16.3	76-561	200	DEX0448_036.orf.3
72	DEX0448_037.nt.1	11p15.5	3-818	201	DEX0448_037.aa.1
73	DEX0448_037.nt.2	11p15.5	3-911	202	DEX0448_037.aa.2
74	DEX0448_037.nt.3	11p15.5	494-1757	203	DEX0448_037.aa.3
74	DEX0448_037.nt.3	11p15.5	1016-1756	204	DEX0448_037.orf.3
75	DEX0448_037.nt.4	11p15.5	240-947	205	DEX0448_037.aa.4
76	DEX0448_037.nt.5	11p15.5	240-1007	206	DEX0448_037.aa.5
77	DEX0448_037.nt.6	11p15.5	240-869	207	DEX0448_037.aa.6
78	DEX0448_038.nt.1	18p11.32	1-179	208	DEX0448_038.aa.1
78	DEX0448_038.nt.1	18p11.32	22-294	209	DEX0448_038.orf.1
79	DEX0448_039.nt.1	11q11	620-877	210	DEX0448_039.aa.1
80	DEX0448_040.nt.1	16p13.3	1-325	211	DEX0448_040.aa.1
80	DEX0448_040.nt.1	16p13.3	52-321	212	DEX0448_040.orf.1
81	DEX0448_040.nt.2	16p13.3	43-625	213	DEX0448_040.aa.2
81	DEX0448_040.nt.2	16p13.3	51-617	214	DEX0448_040.orf.2
82	DEX0448_040.nt.3	16p13.3	22-630	215	DEX0448_040.aa.3
82	DEX0448_040.nt.3	16p13.3	3-626	216	DEX0448_040.orf.3
83	DEX0448_040.nt.4	16p13.3	1-567	217	DEX0448_040.aa.4
84	DEX0448_040.nt.5	16p13.3	1-517	218	DEX0448_040.aa.5
84	DEX0448_040.nt.5	16p13.3	1-513	219	DEX0448_040.orf.5
85	DEX0448_040.nt.6	16p13.3	1352-1823	220	DEX0448_040.aa.6
85	DEX0448_040.nt.6	16p13.3	1347-1814	221	DEX0448_040.orf.6
86	DEX0448_040.nt.7	16p13.3	34-265	222	DEX0448_040.aa.7
86	DEX0448_040.nt.7	16p13.3	504-920	223	DEX0448_040.orf.7
87	DEX0448_041.nt.1	2q37.1	151-1854	224	DEX0448_041.aa.1
88	DEX0448_041.nt.2	2q37.1	151-1710	225	DEX0448_041.aa.2
89	DEX0448_042.nt.1	2q33.1	188-1765	226	DEX0448_042.aa.1
90	DEX0448_043.nt.1	15q22.31	1-366	227	DEX0448_043.aa.1
90	DEX0448_043.nt.1	15q22.31	154-366	228	DEX0448_043.orf.1
91	DEX0448_043.nt.2	15q22.31	1-730	229	DEX0448_043.aa.2

91	DEX0448_043.nt.2	15q22.31	24-1049	230	DEX0448_043.orf.2
92	DEX0448_044.nt.1	20q13.33	1-219	231	DEX0448_044.aa.1
92	DEX0448_044.nt.1	20q13.33	48-356	232	DEX0448_044.orf.1
93	DEX0448_044.nt.2	20q13.33	1-219	231	DEX0448_044.aa.1
93	DEX0448_044.nt.2	20q13.33	48-383	233	DEX0448_044.orf.2
94	DEX0448_044.nt.3	20q13.33	1-265	234	DEX0448_044.aa.3
94	DEX0448_044.nt.3	20q13.33	2-259	235	DEX0448_044.orf.3
95	DEX0448_044.nt.4	20q13.33	31-265	236	DEX0448_044.aa.4
95	DEX0448_044.nt.4	20q13.33	3-260	237	DEX0448_044.orf.4

The polypeptides of the present invention were analyzed and the following attributes were identified; specifically, epitopes, post translational modifications, signal peptides and transmembrane domains. Antigenicity (Epitope) prediction was performed through the antigenic module in the EMBOSS package. Rice, P., EMBOSS: The European Molecular Biology Open Software Suite, *Trends in Genetics* 16(6): 276-277 (2000). The antigenic module predicts potentially antigenic regions of a protein sequence, using the method of Kolaskar and Tongaonkar. Kolaskar, AS and Tongaonkar, PC., A semi-empirical method for prediction of antigenic determinants on protein antigens, *FEBS Letters* 276: 172-174 (1990). Examples of post-translational modifications (PTMs) and other motifs of the CSPs of this invention are listed below. In addition, antibodies that specifically bind such post-translational modifications may be useful as a diagnostic or as therapeutic. The PTMs and other motifs were predicted by using the ProSite Dictionary of Proteins Sites and Patterns (Bairoch *et al.*, *Nucleic Acids Res.* 25(1):217-221 (1997)), the following motifs, including PTMs, were predicted for the CSPs of the invention. The signal peptides were detected by using the SignalP 2.0, *see Nielsen et al.*, *Protein Engineering* 12, 3-9 (1999). Prediction of transmembrane helices in proteins was performed by the application TMHMM 2.0, "currently the best performing transmembrane prediction program", according to authors (Krogh *et al.*, *Journal of Molecular Biology*, 305(3):567-580, (2001); Moller *et al.*, *Bioinformatics*, 17(7):646-653, (2001); Sonnhammer, *et al.*, *A hidden Markov model for predicting transmembrane helices in protein sequences* in Glasgow, *et al.* Ed. Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology, pages 175-182, Menlo Park, CA, 1998. AAAI Press. The PSORT II program may also be used to predict cellular localizations. Horton *et al.*, *Intelligent Systems for Molecular Biology* 5: 147-152 (1997). The table below includes the following sequence annotations: Signal peptide presence;

TM (number of membrane domain, topology in orientation and position); Amino acid location and antigenic index (location, AI score); PTM and other motifs (type, amino acid residue locations); and functional domains (type, amino acid residue locations).

DEX ID	Sig P	TMHMM	Antigenicity	PTM	Domains
DEX0448_001.aa.1	N	0 - 01-185;		CK2_PHOSPHO_SITE 157-160; CAMP_PHOSPHO_SITE 135-138; PKC_PHOSPHO_SITE 104-106; ASN_GLYCOSYLATION 35-38; PKC_PHOSPHO_SITE 178-180; CAMP_PHOSPHO_SITE 101-104; PKC_PHOSPHO_SITE 166-168; PKC_PHOSPHO_SITE 70-72; CK2_PHOSPHO_SITE 155-158; CK2_PHOSPHO_SITE 104-107; PKC_PHOSPHO_SITE 94-96; CK2_PHOSPHO_SITE 70-73; PKC_PHOSPHO_SITE 124-126;	
DEX0448_001.orf.1	N	0 - 01-109;		PKC_PHOSPHO_SITE 56-58; PKC_PHOSPHO_SITE 83-85; MYRISTYL 40-45; CAMP_PHOSPHO_SITE 95-98; CAMP_PHOSPHO_SITE 35-38; PKC_PHOSPHO_SITE 44-46; MYRISTYL 56-61; PKC_PHOSPHO_SITE 30-32; PKC_PHOSPHO_SITE 57-59; AMIDATION 10-13; PKC_PHOSPHO_SITE 11-13; MYRISTYL 26-31; PKC_PHOSPHO_SITE 98-100; AMIDATION 56-59; CK2_PHOSPHO_SITE 22-25; CAMP_PHOSPHO_SITE 106-109; CAMP_PHOSPHO_SITE 36-39; PKC_PHOSPHO_SITE 27-29; CAMP_PHOSPHO_SITE 53-56; PKC_PHOSPHO_SITE 22-24; CK2_PHOSPHO_SITE 38-41;	
DEX0448_002.aa.1	N	0 - 01-106;	79-91,1.14; 37-43,1.089; 69-77,1.12; 17-29,1.266; 46-58,1.199; 95-101,1.108;	PKC_PHOSPHO_SITE 81-83; CK2_PHOSPHO_SITE 77-80; CK2_PHOSPHO_SITE 68-71; PKC_PHOSPHO_SITE 91-93;	
DEX0448_002.orf.1	N	0 - 11-75;	64-70,1.108; 38-46,1.12; 21-27,1.092; 9-16,1.089; 48-60,1.14;	PKC_PHOSPHO_SITE 20-22; CK2_PHOSPHO_SITE 6-9; PKC_PHOSPHO_SITE 6-8; CK2_PHOSPHO_SITE 37-40; PKC_PHOSPHO_SITE 50-52; CK2_PHOSPHO_SITE 46-49; PKC_PHOSPHO_SITE 60-62; TYR PHOSPHO SITE	

DEX0448_003.aa.1	N	1 - il-87;tm88-110;ol11-224;	178-184,1.074; 39-45,1.078; 4-11,1.127; 26-33,1.094; 156-165,1.11; 202-217,1.171; 53-66,1.112; 116-141,1.142; 85-109,1.211; 68-80,1.18;	5-13; PKC_PHOSPHO_SITE 169-171; PKC_PHOSPHO_SITE 155-157; CK2_PHOSPHO_SITE 220-223; CK2_PHOSPHO_SITE 44-47; TYR_PHOSPHO_SITE 160-166; CK2_PHOSPHO_SITE 186-189; MYRISTYL 112-117; CK2_PHOSPHO_SITE 111-114; MYRISTYL 52-57; PKC_PHOSPHO_SITE 44-46; MYRISTYL 55-60; PKC_PHOSPHO_SITE 67-69; PKC_PHOSPHO_SITE 141-143; MYRISTYL 61-66; MYRISTYL 9-14; PKC_PHOSPHO_SITE 186-188; CK2_PHOSPHO_SITE 184-187; CAMP_PHOSPHO_SITE 84-87;	SP_P07858_CATB_HUMA N 29-199; THIOL_PROTEASE_CYS 51-62; Peptidase_C1 29-223; Pept_C1 29-223;
DEX0448_003.orf.1	Y	0 - ol-181;		PKC_PHOSPHO_SITE 30-32; MICROBODIES_CTER 179-181; CK2_PHOSPHO_SITE 21-24; MYRISTYL 32-37; MYRISTYL 34-39; CAMP_PHOSPHO_SITE 67-70; MYRISTYL 163-168; MYRISTYL 29-34; ASN_GLYCOSYLATION 34-37; MYRISTYL 37-42; ASN_GLYCOSYLATION 19-22; PKC_PHOSPHO_SITE 126-128; PKC_PHOSPHO_SITE 149-151; ASN_GLYCOSYLATION 28-31; MYRISTYL 153-158; MYRISTYL 134-139; CK2_PHOSPHO_SITE 126-129; ASN_GLYCOSYLATION 69-72; MYRISTYL 166-171; MYRISTYL 137-142; ASN_GLYCOSYLATION 31-34; PKC_PHOSPHO_SITE 21-23; MYRISTYL 91-96; MYRISTYL 26-31; MYRISTYL 143-148;	THIOL_PROTEASE_CYS 133-144; SP_P07858_CATB_HUMA N 111-160;
DEX0448_004.aa.1	N	0 - ol-227;		PKC_PHOSPHO_SITE 30-32; MYRISTYL 48-53; PKC_PHOSPHO_SITE 33-35; CK2_PHOSPHO_SITE 75-78; CK2_PHOSPHO_SITE 202-205; AMIDATION 176-179; ASN_GLYCOSYLATION 85-88; CK2_PHOSPHO_SITE 116-119; PKC_PHOSPHO_SITE 59-61; AMIDATION 83-86; MYRISTYL 45-50; ASN_GLYCOSYLATION 83-86; CK2_PHOSPHO_SITE 149-152; PKC_PHOSPHO_SITE 67-69; MYRISTYL 78-83; TYR_PHOSPHO_SITE 186-194; CK2_PHOSPHO_SITE 2-5; TYR_PHOSPHO_SITE 70-	

				78;	
DEX0448_004.orf.1	N	0 - 01-222;		CK2_PHOSPHO_SITE 98-101; PKC_PHOSPHO_SITE 64-66; MYRISTYL 201-206; MYRISTYL 149-154; PKC_PHOSPHO_SITE 215-217; CK2_PHOSPHO_SITE 59-62; AMIDATION 43-46; CK2_PHOSPHO_SITE 7-10; PKC_PHOSPHO_SITE 4-6; MYRISTYL 44-49; MYRISTYL 181-186; MYRISTYL 23-28; MYRISTYL 12-17; TYR_PHOSPHO_SITE 151-157; ASN_GLYCOSYLATION 148-151; PKC_PHOSPHO_SITE 19-21; MYRISTYL 185-190;	
DEX0448_005.aa.1	Y	0 - 01-74;	4-15,1.203; 62-71,1.149; 18-39,1.368;	MYRISTYL 33-38; RGD 16-18; MYRISTYL 46-51; CK2_PHOSPHO_SITE 49-52;	
DEX0448_005.orf.1	Y	0 - 01-85;	29-50,1.368; 19-26,1.159; 73-82,1.149; 4-11,1.223;	CK2_PHOSPHO_SITE 60-63; MYRISTYL 44-49; MYRISTYL 57-62; CK2_PHOSPHO_SITE 16-19; MYRISTYL 1-6; RGD 27-29; PKC_PHOSPHO_SITE 16-18; MYRISTYL 5-10;	
DEX0448_005.orf.2	Y	0 - 01-85;	4-11,1.223; 19-26,1.159; 73-82,1.149; 29-50,1.368;	CK2_PHOSPHO_SITE 60-63; MYRISTYL 44-49; MYRISTYL 57-62; MYRISTYL 1-6; RGD 27-29; PKC_PHOSPHO_SITE 16-18; MYRISTYL 5-10;	
DEX0448_006.aa.1	N	0 - 01-66;	4-13,1.116; 26-40,1.141; 50-63,1.077;	PKC_PHOSPHO_SITE 16-18; CK2_PHOSPHO_SITE 6-9; CK2_PHOSPHO_SITE 55-58; CK2_PHOSPHO_SITE 40-43; PKC_PHOSPHO_SITE 64-66; PKC_PHOSPHO_SITE 1-3;	
DEX0448_006.orf.1	N	0 - 01-66;		PKC_PHOSPHO_SITE 1-3; PKC_PHOSPHO_SITE 41-43; CK2_PHOSPHO_SITE 6-9; PKC_PHOSPHO_SITE 16-18; ASN_GLYCOSYLATION 54-57; PKC_PHOSPHO_SITE 45-47; ASN_GLYCOSYLATION 52-55; MYRISTYL 54-59;	
DEX0448_007.aa.1	N	0 - 01-126;	50-61,1.121; 95-108,1.164; 110-115,1.07; 64-90,1.164; 4-30,1.198; 35-	ASN_GLYCOSYLATION 75-78; MYRISTYL 103-108; CK2_PHOSPHO_SITE 116-119; MYRISTYL 95-100; CK2_PHOSPHO_SITE 111-114; ASN_GLYCOSYLATION 109-112; MYRISTYL 97-102;	

136

				44, 1.091;			
DEX0448_007.orf.1	Y	0 - 01-117;		55-64, 1.091; 70-81, 1.121; 84-110, 1.164; 4-50, 1.198;	ASN GLYCOSYLATION 95-98;		
DEX0448_008.aa.1	N	0 - 01-170;		144-149, 1.042; 118-125, 1.083; 16-29, 1.194; 68-74, 1.049; 35-57, 1.174; 90-105, 1.147; 153-159, 1.055; 132-141, 1.057; 107-116, 1.088; 161-167, 1.227;	PKC_PHOSPHO_SITE 143-145; PKC_PHOSPHO_SITE 148-150; MYRISTYL 43-48; CAMP_PHOSPHO_SITE 31-34;		
DEX0448_008.orf.1	N	0 - 01-225;			CK2_PHOSPHO_SITE 28-31; MYRISTYL 101-106; PKC_PHOSPHO_SITE 62-64; PKC_PHOSPHO_SITE 55-57; PKC_PHOSPHO_SITE 108-110; CK2_PHOSPHO_SITE 98-101; PKC_PHOSPHO_SITE 203-205; MYRISTYL 48-53; MYRISTYL 95-100; PKC_PHOSPHO_SITE 57-59; PKC_PHOSPHO_SITE 198-200; PKC_PHOSPHO_SITE 61-63; MYRISTYL 113-118; CK2_PHOSPHO_SITE 41-44; ASN_GLYCOSYLATION 68-71; MYRISTYL 41-46; CK2_PHOSPHO_SITE 40-43; MYRISTYL 110-115; MYRISTYL 57-62; PKC_PHOSPHO_SITE 40-42;		
DEX0448_009.aa.1	N	0 - 01-175;		51-65, 1.157; 123-132, 1.171; 22-41, 1.188; 161-172, 1.147; 147-156, 1.167; 80-111, 1.118; 9-15, 1.083;	CK2_PHOSPHO_SITE 86-89; MYRISTYL 43-48; MYRISTYL 72-77; PKC_PHOSPHO_SITE 159-161; MYRISTYL 138-143; MYRISTYL 1-6; MYRISTYL 154-159;		
DEX0448_010.aa.1	N	1 - 11-32; tm33-55; 056-270;		51-59, 1.108; 190-216, 1.143; 64-72, 1.071;	MYRISTYL 47-52; AMIDATION 17-20; PKC_PHOSPHO_SITE 241-243; ASN GLYCOSYLATION 120-123; MYRISTYL 230-		

			223-231, 1.133; 135-145, 1.157; 125-133, 1.133; 234-240, 1.053; 29-48, 1.206; 95-101, 1.103; 82-89, 1.071; 253-259, 1.076; 170-177, 1.078;	235; ASN GLYCOSYLATION 114-117; CK2 PHOSPHO_SITE 9-12; MYRISTYL 31-36; MYRISTYL 206-211; MYRISTYL 22-27; CK2 PHOSPHO_SITE 145-148; MYRISTYL 252-257;	
DEX0448_011.aa.1	N	0 - 01-225;	62-72, 1.069; 5-17, 1.236; 52-57, 1.071; 198-210, 1.111; 110-121, 1.105; 87-96, 1.158; 21-38, 1.084; 130-140, 1.146; 178-191, 1.123;	PKC PHOSPHO_SITE 65-67; CK2 PHOSPHO_SITE 162-165; MYRISTYL 172-177; ASN GLYCOSYLATION 198-201; MYRISTYL 175-180; CK2 PHOSPHO_SITE 89-92; PKC PHOSPHO_SITE 130-132; PKC PHOSPHO_SITE 82-84; TYR PHOSPHO_SITE 39-46;	sp_076045_076045_HU MAN 76-225; COLFI 76-225; COLFI 59-225;
DEX0448_011.orf.1	N	0 - 01-121;		MYRISTYL 71-76; MYRISTYL 68-73; PKC PHOSPHO_SITE 26-28; CK2 PHOSPHO_SITE 58-61; ASN GLYCOSYLATION 94-97;	sp_076045_076045_HU MAN 1-121; COLFI 3-121; COLFI 6-121;
DEX0448_012.aa.1	Y	1 - i1-6; tm7-29; o30-66;	8-55, 1.195;	TYR PHOSPHO_SITE 55-63;	
DEX0448_012.orf.1	N	0 - i1-81;		PKC PHOSPHO_SITE 53-55; CAMP PHOSPHO_SITE 55-58;	
DEX0448_013.aa.1	N	0 - 01-253;	76-81, 1.064; 4-19, 1.152; 166-183, 1.13; 22-33, 1.089; 49-61, 1.108; 112-127, 1.133; 147-157, 1.104; 229-244, 1.17;	PKC PHOSPHO_SITE 20-22; MYRISTYL 187-192; PKC PHOSPHO_SITE 209-211; PKC PHOSPHO_SITE 199-201; PKC PHOSPHO_SITE 46-48; PKC PHOSPHO_SITE 45-47; PKC PHOSPHO_SITE 96-98; MYRISTYL 191-196; PKC PHOSPHO_SITE 210-212; PKC PHOSPHO_SITE 119-121; CK2 PHOSPHO_SITE 160-163; CK2 PHOSPHO_SITE 89-92; CK2 PHOSPHO_SITE 85-88;	DnaJ 2-64; DnaJ_1 49-68; DnaJ_2 3-72; DnaJ 3-72;
DEX0448_013.orf.1	N	0 - 01-203;		PKC PHOSPHO_SITE 4-6; AMIDATION 30-33; MYRISTYL 101-106; CAMP PHOSPHO_SITE 1-4;	DnaJ 1-31; DnaJ_1 8-27; DnaJ_2 1-31;

					PKC_PHOSPHO_SITE 168-170; CK2_PHOSPHO_SITE 40-43; PKC_PHOSPHO_SITE 167-169; PKC_PHOSPHO_SITE 32-34; AMIDATION 31-34; MYRISTYL 104-109; PKC_PHOSPHO_SITE 157-159; MYRISTYL 145-150; PKC_PHOSPHO_SITE 102-104; TYR_PHOSPHO_SITE 24-32; MYRISTYL 64-69; MYRISTYL 149-154; PKC_PHOSPHO_SITE 5-7; CK2_PHOSPHO_SITE 118-121;	
DEX0448_014.aa.1	N	0 - i1-128;	34-42,1.121; 69-76,1.133; 81-91,1.1; 112-123,1.129; 9-16,1.074; 48-53,1.079; 93-110,1.126;		MYRISTYL 15-20; PKC_PHOSPHO_SITE 109-111; AMIDATION 62-65; TYR_PHOSPHO_SITE 27-34; PKC_PHOSPHO_SITE 6-8; CK2_PHOSPHO_SITE 44-47; CK2_PHOSPHO_SITE 35-38;	KH_TYPE_2 21-92; KH 47-95;
DEX0448_014.orf.1	N	0 - o1-143;			ASN GLYCOSYLATION 126-129; PKC_PHOSPHO_SITE 128-130; MYRISTYL 13-18; MYRISTYL 123-128; MYRISTYL 5-10; PKC_PHOSPHO_SITE 21-23; MYRISTYL 30-35; ASN GLYCOSYLATION 2-5; TYR_PHOSPHO_SITE 42-49; CK2_PHOSPHO_SITE 59-62; CAMP_PHOSPHO_SITE 1-4; CK2_PHOSPHO_SITE 50-53; MYRISTYL 2-7; CK2_PHOSPHO_SITE 115-118; AMIDATION 77-80;	KH_TYPE_2 36-107; KH 62-110;
DEX0448_015.aa.1	N	0 - o1-75;			CAMP_PHOSPHO_SITE 24-27; MYRISTYL 23-28; MYRISTYL 9-14; PKC_PHOSPHO_SITE 27-29; AMIDATION 27-30; PKC_PHOSPHO_SITE 23-25; CK2_PHOSPHO_SITE 40-43; ASN GLYCOSYLATION 19-22; MYRISTYL 11-16; MYRISTYL 61-66;	
DEX0448_015.orf.1	N	0 - o1-110;	17-26,1.082; 101-107,1.066; 6-13,1.147;		PKC_PHOSPHO_SITE 3-5; AMIDATION 35-38; MYRISTYL 47-52; PKC_PHOSPHO_SITE 26-28; MYRISTYL 66-71; MYRISTYL 2-7; MYRISTYL 13-18; PKC_PHOSPHO_SITE 70-72;	
DEX0448_016.aa.1	N	0 - o1-196;	92-105,1.124; 146-160,1.144;		CK2_PHOSPHO_SITE 21-24; MYRISTYL 17-22; CK2_PHOSPHO_SITE 150-153; PKC_PHOSPHO_SITE 104-106; MYRISTYL 173-178;	RIBOSOMAL_SITE 102-117: Ribosomal Site

139

			70-84, 1.113; 8-37, 1.201; 175-183, 1.1; 167-173, 1.085; 52-68, 1.147; 127-134, 1.073; 112-120, 1.107; 42-50, 1.126;	PKC PHOSPHO_SITE 131-133; CK2_PHOSPHO_SITE 17-20; PKC_PHOSPHO_SITE 91-93; TYR_PHOSPHO_SITE 75-82; PKC_PHOSPHO_SITE 26-28; PKC_PHOSPHO_SITE 2-4; PKC_PHOSPHO_SITE 191-193; PKC_PHOSPHO_SITE 69-71;	63-183;
DEX0448_016.aa.2	N	0 - 01-207;	4-19, 1.128; 46-52, 1.028; 123-131, 1.107; 81-95, 1.113; 157-171, 1.144; 103-116, 1.124; 186-194, 1.1; 63-69, 1.06; 138-145, 1.073; 178-184, 1.085;	TYR_PHOSPHO_SITE 86-93; PKC_PHOSPHO_SITE 80-82; MYRISTYL 18-23; CK2_PHOSPHO_SITE 161-164; PKC_PHOSPHO_SITE 115-117; MYRISTYL 184-189; MYRISTYL 48-53; PKC_PHOSPHO_SITE 142-144; PKC_PHOSPHO_SITE 202-204; PKC_PHOSPHO_SITE 102-104;	Ribosomal_S17e 73-194; RIBOSOMAL_S17E 113-128;
DEX0448_016.orf.2	N	0 - 01-180;	159-167, 1.1; 130-144, 1.144; 151-157, 1.085; 96-104, 1.107; 28-33, 1.028; 4-15, 1.09; 111-118, 1.073; 76-89, 1.124; 54-68, 1.113;	PKC_PHOSPHO_SITE 53-55; MYRISTYL 157-162; MYRISTYL 43-48; MYRISTYL 1-6; PKC_PHOSPHO_SITE 175-177; PKC_PHOSPHO_SITE 75-77; PKC_PHOSPHO_SITE 88-90; TYR_PHOSPHO_SITE 59-66; PKC_PHOSPHO_SITE 115-117; MYRISTYL 29-34; CK2_PHOSPHO_SITE 134-137;	Ribosomal_S17e 46-167; RIBOSOMAL_S17E 86-101;
DEX0448_016.aa.3	N	0 - 01-150;	114-134, 1.13; 31-44, 1.124; 66-73, 1.073; 9-23, 1.113; 85-99, 1.144; 106-112, 1.085; 51-59, 1.107;	CK2_PHOSPHO_SITE 89-92; PKC_PHOSPHO_SITE 43-45; PKC_PHOSPHO_SITE 30-32; PKC_PHOSPHO_SITE 145-147; PKC_PHOSPHO_SITE 8-10; MYRISTYL 112-117; PKC_PHOSPHO_SITE 70-72; TYR_PHOSPHO_SITE 14-21;	Ribosomal_S17e 1-122; RIBOSOMAL_S17E 41-56;
DEX0448_017.aa.1	Y	0 - 01-298;	139-146, 1.126; 53-64, 1.13;	MYRISTYL 172-177; MYRISTYL 48-53; MYRISTYL 151-156; AMIDATION 131-134;	Ribosomal_L2_C 137-272; Ribosomal L2

			237-243,1.073; 245-251,1.062; 125-134,1.124; 203-213,1.219; 115-122,1.054; 218-230,1.059; 188-201,1.075; 82-92,1.106; 151-159,1.199; 96-107,1.166; 273-281,1.102; 69-76,1.084; 174-180,1.112; 4-21,1.195;	PKC_PHOSPHO_SITE 284-286; PKC_PHOSPHO_SITE 195-197; MYRISTYL 122-127; CK2_PHOSPHO_SITE 147-150; AMIDATION 272-275; PKC_PHOSPHO_SITE 184-186; MYRISTYL 208-213; MYRISTYL 96-101; PKC_PHOSPHO_SITE 265-267; CK2_PHOSPHO_SITE 292-295; MYRISTYL 277-282; CAMP_PHOSPHO_SITE 111-114;	52-131; RIBOSOMAL_L2 238-249;
DEX0448_017.orf.1	N	0 - 01-271;		MYRISTYL 21-26; CAMP_PHOSPHO_SITE 84-87; MYRISTYL 2-7; AMIDATION 245-248; MYRISTYL 124-129; MYRISTYL 95-100; PKC_PHOSPHO_SITE 257-259; PKC_PHOSPHO_SITE 157-159; MYRISTYL 1-6; CK2_PHOSPHO_SITE 120-123; PKC_PHOSPHO_SITE 168-170; MYRISTYL 250-255; MYRISTYL 181-186; MYRISTYL 69-74; AMIDATION 104-107; CK2_PHOSPHO_SITE 265-268; PKC_PHOSPHO_SITE 238-240; MYRISTYL 145-150;	Ribosomal_L2_C 110-245; RIBOSOMAL_L2 211-222; Ribosomal_L2 25-104;
DEX0448_018.aa.1	N	0 - 01-550;	244-258,1.163; 279-285,1.083; 160-183,1.089; 316-324,1.167; 306-312,1.088; 8-28,1.104; 191-215,1.121; 449-466,1.21; 134-140,1.089; 109-120,1.138; 290-301,1.11; 359-367,1.08;	CK2_PHOSPHO_SITE 260-263; TYR_PHOSPHO_SITE 303-310; ASN_GLYCOSYLATION 412-415; PKC_PHOSPHO_SITE 435-437; MYRISTYL 530-535; PKC_PHOSPHO_SITE 475-477; MYRISTYL 526-531; PKC_PHOSPHO_SITE 277-279; PKC_PHOSPHO_SITE 78-80; MYRISTYL 173-178; CK2_PHOSPHO_SITE 159-162; PKC_PHOSPHO_SITE 230-232; PKC_PHOSPHO_SITE 29-31; ASN_GLYCOSYLATION 183-186; CK2_PHOSPHO_SITE 210-213; CK2_PHOSPHO_SITE 467-470; CK2_PHOSPHO_SITE 266-269; PKC_PHOSPHO_SITE 379-381; CK2 PHOSPHO SITE	HSP90 5-522; GLN_RICH 8-66;

			267-273, 1.097; 370-393, 1.221; 32-39, 1.056; 122-132, 1.11; 329-337, 1.126; 522-544, 1.199; 148-158, 1.145; 429-436, 1.102; 51-72, 1.095; 481-487, 1.069; 499-509, 1.063;	91-94;		
DEX0448_018.orf.1	N	0 - 01-190;		CK2 PHOSPHO SITE 10-13; ASN GLYCOSYLATION 162-165; PKC PHOSPHO SITE 129-131; CK2 PHOSPHO SITE 16-19; TYR PHOSPHO SITE 53-60; PKC PHOSPHO SITE 27-29;	HSP90 1-189;	
DEX0448_019.aa.1	N	0 - 11-111;	89-95, 1.136; 21-35, 1.188; 79-85, 1.164; 44-51, 1.166;	CAMP PHOSPHO SITE 106-109; PKC PHOSPHO SITE 104-106; ASN GLYCOSYLATION 40-43; CK2 PHOSPHO SITE 85-88; PKC PHOSPHO SITE 56-58; ASN GLYCOSYLATION 77-80;	RIBOSOMAL_L18E 49-66; Ribosomal_L18e 10-111; sp_Q07020_RL18_HUMA N 28-97;	
DEX0448_020.aa.1	N	0 - 01-261;	110-140, 1.126; 207-223, 1.108; 69-93, 1.241; 7-12, 1.048; 147-158, 1.092; 42-48, 1.091; 16-31, 1.163; 176-187, 1.085; 226-246, 1.099; 164-171, 1.122;	MYRISTYL 232-237; CK2 PHOSPHO SITE 245-248; ASN GLYCOSYLATION 254-257; MYRISTYL 66-71; CK2 PHOSPHO SITE 192-195; MYRISTYL 16-21; MYRISTYL 37-42; CK2 PHOSPHO SITE 183-186; MYRISTYL 22-27; CK2 PHOSPHO SITE 80-83; CK2 PHOSPHO SITE 106-109; MYRISTYL 70-75; CK2 PHOSPHO SITE 96-99; PKC PHOSPHO SITE 164-166; CK2 PHOSPHO SITE 224-227; MYRISTYL 32-37;	PROTEIN_KINASE_DOM 1-159; sp_P54754_EPB3_MOUSE 79-161; TYRKINASE 82-104; TyKc 13-155; SAM 186-250; TYRKINASE 126-148; SAM 185-252; SAM_DOMAIN 188-252;	
DEX0448_021.aa.1	N	1 - 11-117; tm118-140; 0141-361;	265-285, 1.106; 116-152, 1.242; 180-188, 1.107; 293-313, 1.064; 4-113, 1.239; 195-208, 1.061;	MYRISTYL 25-30; ASN GLYCOSYLATION 86-89; MYRISTYL 84-89; CK2 PHOSPHO SITE 88-91; ASN GLYCOSYLATION 102-105; LEUCINE_ZIPPER 123-144; PKC PHOSPHO SITE 356-358; MYRISTYL 333-338; MYRISTYL 294-299; PKC PHOSPHO SITE 241-243;	INTRLNK1R1F 123-151; INTRLNK1R1F 178-202; TIR 163-243; PRO_RICH 253-304; INTRLNK1R1F 160-174;	

			332-337, 1.052; 211-258, 1.189; 164-174, 1.144; 315-324, 1.062;	ASN_GLYCOSYLATION 73-76; CK2_PHOSPHO_SITE 170-173; ASN_GLYCOSYLATION 217-220; MYRISTYL 94-99; AMIDATION 338-341; CK2_PHOSPHO_SITE 63-66; MYRISTYL 53-58; ASN_GLYCOSYLATION 31-34; MYRISTYL 59-64; CK2_PHOSPHO_SITE 13-16;	INTRLNIRIF 203- 230; IG_LIKE 9-109;
DEX0448_021.aa.2	N	1 - ol- 117;tm118- 140;i141-329;	211-258, 1.189; 293-313, 1.064; 164-174, 1.144; 265-285, 1.106; 180-188, 1.107; 4-113, 1.239; 195-208, 1.061; 116-152, 1.242;	CK2_PHOSPHO_SITE 13-16; ASN_GLYCOSYLATION 86-89; CK2_PHOSPHO_SITE 170-173; MYRISTYL 59-64; CK2_PHOSPHO_SITE 63-66; MYRISTYL 316-321; ASN_GLYCOSYLATION 217-220; MYRISTYL 294-299; PKC_PHOSPHO_SITE 241- 243; CK2_PHOSPHO_SITE 88-91; MYRISTYL 53- 58; ASN_GLYCOSYLATION 73-76; LEUCINE_ZIPPER 123-144; MYRISTYL 94-99; MYRISTYL 84-89; ASN_GLYCOSYLATION 102-105; ASN_GLYCOSYLATION 31-34; MYRISTYL 25-30;	INTRLNIRIF 178- 202; INTRLNIRIF 160-174; TIR 163- 243; INTRLNIRIF 123-151; PRO_RICH 253-304; INTRLNIRIF 203- 230; IG_LIKE 9-109;
DEX0448_021.aa.3	N	0 - ol-164;	85-91, 1.094; 154-161, 1.177; 38-50, 1.149; 21-28, 1.047; 112-120, 1.1; 122-132, 1.078; 54-70, 1.128; 137-144, 1.091;	PKC_PHOSPHO_SITE 150-152; PKC_PHOSPHO_SITE 11-13; CK2_PHOSPHO_SITE 137-140; CK2_PHOSPHO_SITE 160-163; MYRISTYL 128- 133; CK2_PHOSPHO_SITE 100-103; ASN_GLYCOSYLATION 148-151; CK2_PHOSPHO_SITE 130-133;	
DEX0448_022.aa.1	N	1 - il- 27;tm28- 50;o51-66;	29-50, 1.193;	CK2_PHOSPHO_SITE 7-10; MYRISTYL 58-63;	
DEX0448_022.orf.1	N	0 - ol-135;	4-15, 1.159; 18- 28, 1.073; 105- 112, 1.072; 41- 100, 1.205; 30- 37, 1.12;	MYRISTYL 37-42; PKC_PHOSPHO_SITE 3-5; CK2_PHOSPHO_SITE 99-102;	
DEX0448_022.aa.2	N	0 - ol-115;	5-15, 1.073; 17- 88, 1.205; 93- 100, 1.072;	CK2_PHOSPHO_SITE 87-90;	

DEX0448_022.orf.2	N	0 - 01-135;	105-112,1.072; 41-100,1.205; 30-37,1.12; 18- 28,1.073; 4- 15,1.159;	MYRISTYL 37-42; PKC_PHOSPHO_SITE 3-5; CK2_PHOSPHO_SITE 99-102;	
DEX0448_023.aa.1	N	0 - 01-220;	87-95,1.12; 182-202,1.117; 173-180,1.104; 13-51,1.216; 146-151,1.03; 99-128,1.174; 74-79,1.072; 157-166,1.153;	CK2_PHOSPHO_SITE 67-70; MYRISTYL 212-217; CK2_PHOSPHO_SITE 139-142; CK2_PHOSPHO_SITE 153-156; MYRISTYL 214-219; CK2_PHOSPHO_SITE 8-11; MYRISTYL 168-173; MYRISTYL 186-191;	G6PISOMERASE 39-60; G6PISOMERASE 168- 186; G6PISOMERASE 186-200; PGI 1-220; G6PISOMERASE 200- 213;
DEX0448_023.orf.1	N	0 - 01-287;		MYRISTYL 3-8; CK2_PHOSPHO_SITE 265-268; MYRISTYL 243-248; MYRISTYL 217-222; CK2_PHOSPHO_SITE 184-187; CK2_PHOSPHO_SITE 39-42; MYRISTYL 199-204; CK2_PHOSPHO_SITE 98-101; CK2_PHOSPHO_SITE 170-173; MYRISTYL 10-15; MYRISTYL 6-11;	G6PISOMERASE 199- 217; G6PISOMERASE 70-91; PGI 2-275; G6PISOMERASE 231- 244; P_GLUCOSE_ISOMERASE 2 231-248; G6PISOMERASE 217- 231;
DEX0448_023.aa.2	Y	0 - 01-147;	9-39,1.137; 121-127,1.071; 62-85,1.145; 42-60,1.157; 101-109,1.079;	MYRISTYL 103-108; PKC_PHOSPHO_SITE 84-86; MYRISTYL 6-11; CK2_PHOSPHO_SITE 44-47; CK2_PHOSPHO_SITE 125-128; CK2_PHOSPHO_SITE 74-77; MYRISTYL 94-99; PKC_PHOSPHO_SITE 88-90;	
DEX0448_023.orf.2	N	0 - 01-150;		CK2_PHOSPHO_SITE 110-113; CK2_PHOSPHO_SITE 146-149; CK2_PHOSPHO_SITE 102-105; MYRISTYL 98-103; CK2_PHOSPHO_SITE 76-79; MYRISTYL 58-63; CK2_PHOSPHO_SITE 136-139; MYRISTYL 90-95; CK2_PHOSPHO_SITE 82-85;	
DEX0448_023.aa.3	N	0 - 01-811;	568-574,1.073; 604-610,1.105; 409-415,1.106; 521-527,1.098;	PKC_PHOSPHO_SITE 113-115; CK2_PHOSPHO_SITE 501-504; MYRISTYL 552-557; MYRISTYL 677- 682; MYRISTYL 136-141; PKC_PHOSPHO_SITE 707-709; CK2 PHOSPHO_SITE 659-662;	GLY_RICH 22-689;

			<p>298-304,1.112; 101-106,1.071; 719-724,1.033; 114-120,1.079; 447-454,1.12; 175-180,1.071; 150-156,1.112; 4-10,1.083; 691-696,1.085; 334-341,1.141; 667-673,1.048; 748-762,1.127; 39-45,1.079; 372-378,1.085; 768-773,1.036; 433-441,1.103; 642-647,1.079; 64-69,1.071; 360-365,1.079; 617-624,1.112; 261-268,1.112; 249-259,1.102; 187-195,1.079; 275-280,1.033; 224-232,1.079; 791-798,1.112; 78-83,1.072; 386-391,1.033;</p>	<p>MYRISTYL 247-252; MYRISTYL 740-745; MYRISTYL 62-67; MYRISTYL 25-30; MYRISTYL 331-336; PKC_PHOSPHO_SITE 611-613; MYRISTYL 649-654; PKC_PHOSPHO_SITE 76-78; MYRISTYL 99-104; MYRISTYL 564-569; MYRISTYL 173-178; AMIDATION 403-406; PKC_PHOSPHO_SITE 187-189; PKC_PHOSPHO_SITE 39-41; MYRISTYL 543-548; PKC_PHOSPHO_SITE 520-522; MYRISTYL 728-733; PKC_PHOSPHO_SITE 403-405; MYRISTYL 555-560; MYRISTYL 506-511; MYRISTYL 652-657; MYRISTYL 210-215; PKC_PHOSPHO_SITE 224-226; CK2_PHOSPHO_SITE 538-541;</p>	
DEX0448_023.orf.3	N	0 - 01-442;	<p>301-307,1.112; 389-394,1.033; 375-381,1.085; 228-235,1.079; 264-271,1.112; 81-86,1.072; 278-283,1.033; 412-418,1.106; 4-12,1.083;</p>	<p>MYRISTYL 28-33; MYRISTYL 65-70; AMIDATION 406-409; MYRISTYL 213-218; PKC_PHOSPHO_SITE 227-229; MYRISTYL 139-144; MYRISTYL 250-255; MYRISTYL 334-339; PKC_PHOSPHO_SITE 406-408; PKC_PHOSPHO_SITE 190-192; PKC_PHOSPHO_SITE 116-118; MYRISTYL 176-181; MYRISTYL 102-107; PKC_PHOSPHO_SITE 79-81; PKC_PHOSPHO_SITE 42-44;</p>	

				191-198,1.079; 42-48,1.079; 153-159,1.112; 252-262,1.102; 117-123,1.079; 337-344,1.141;			
DEX0448_024.aa.1	N	0 - 01-351;		103-112,1.117; 9-31,1.177; 304-309,1.077; 217-222,1.045; 59-78,1.197; 207-212,1.07; 134-145,1.093; 147-162,1.141; 120-132,1.067; 34-51,1.124; 174-186,1.127; 194-204,1.129;		PKC_PHOSPHO_SITE 252-254; PKC_PHOSPHO_SITE 19-21; PKC_PHOSPHO_SITE 280-282; CK2_PHOSPHO_SITE 57-60; TYR_PHOSPHO_SITE 56-63; PKC_PHOSPHO_SITE 168-170; MYRISTYL 185-190; PKC_PHOSPHO_SITE 314-316; PKC_PHOSPHO_SITE 6-8; CK2_PHOSPHO_SITE 6-9; AMIDATION 330-333; PKC_PHOSPHO_SITE 169-171; PKC_PHOSPHO_SITE 263-265; MYRISTYL 240-245; PKC_PHOSPHO_SITE 262-264; CAMP_PHOSPHO_SITE 55-58;	Ribosomal_L18ae 6-176;
DEX0448_024.orf.1	N	0 - 01-223;				TYR_PHOSPHO_SITE 68-75; PKC_PHOSPHO_SITE 181-183; CK2_PHOSPHO_SITE 4-7; PKC_PHOSPHO_SITE 180-182; MYRISTYL 5-10; CK2_PHOSPHO_SITE 18-21; PKC_PHOSPHO_SITE 31-33; PKC_PHOSPHO_SITE 18-20; MYRISTYL 197-202; CK2_PHOSPHO_SITE 69-72; CAMP_PHOSPHO_SITE 67-70;	Ribosomal_L18ae 18-188;
DEX0448_025.aa.1	N	0 - 01-260;		31-39,1.092; 149-156,1.124; 214-229,1.133; 4-25,1.135; 88-98,1.102; 243-257,1.196; 174-190,1.137; 116-122,1.102; 198-205,1.096; 71-81,1.125; 55-60,1.028;		CK2_PHOSPHO_SITE 96-99; ASN GLYCOSYLATION 185-188; MYRISTYL 94-99; MYRISTYL 61-66; MYRISTYL 130-135; MYRISTYL 26-31; PKC_PHOSPHO_SITE 163-165; MYRISTYL 157-162; AMIDATION 193-196; MYRISTYL 33-38; MYRISTYL 29-34; MYRISTYL 79-84; MYRISTYL 159-164; PKC_PHOSPHO_SITE 14-16; MYRISTYL 37-42; PKC_PHOSPHO_SITE 66-68; CAMP_PHOSPHO_SITE 146-149; MYRISTYL 62-67; CK2_PHOSPHO_SITE 42-45;	GLY_RICH 21-142;

146

DEX0448_025.orf.1	N	0 - 01-259;	115-121,1.102; 213-228,1.133; 4-24,1.135; 30- 38,1.092; 197- 204,1.096; 173- 189,1.137; 54- 59,1.028; 148- 155,1.124; 87- 97,1.102; 242- 256,1.196; 70- 80,1.125;	CK2_PHOSPHO_SITE 41-44; PKC_PHOSPHO_SITE 65-67; CK2_PHOSPHO_SITE 95-98; MYRISTYL 156-161; MYRISTYL 78-83; MYRISTYL 28-33; MYRISTYL 60-65; MYRISTYL 32-37; ASN GLYCOSYLATION 184-187; PKC_PHOSPHO_SITE 13-15; MYRISTYL 93-98; MYRISTYL 36-41; MYRISTYL 25-30; MYRISTYL 129-134; AMIDATION 192-195; PKC_PHOSPHO_SITE 162-164; MYRISTYL 158-163; CAMP_PHOSPHO_SITE 145-148; MYRISTYL 61-66;	GLY_RICH 20-141;
DEX0448_026.aa.1	N	1 - 01-259;tm260-282;i283-650;	233-242,1.159; 370-378,1.142; 316-342,1.129; 138-148,1.099; 259-312,1.326; 18-34,1.089; 213-225,1.271; 381-390,1.102; 491-503,1.078; 615-624,1.092; 402-409,1.095; 183-189,1.086; 47-55,1.089; 161-178,1.17; 245-255,1.105; 57-66,1.062; 120-135,1.106; 449-454,1.054; 69-114,1.233; 150-155,1.069; 639-647,1.109;	PKC_PHOSPHO_SITE 401-403; MYRISTYL 278-283; TYR_PHOSPHO_SITE 530-536; MYRISTYL 5-10; MYRISTYL 517-522; CK2_PHOSPHO_SITE 185-188; TYR_PHOSPHO_SITE 546-552; MYRISTYL 482-487; CK2_PHOSPHO_SITE 222-225; PKC_PHOSPHO_SITE 424-426; MYRISTYL 13-18; MYRISTYL 181-186; PKC_PHOSPHO_SITE 172-174; CK2_PHOSPHO_SITE 436-439; CK2_PHOSPHO_SITE 504-507; PKC_PHOSPHO_SITE 467-469; PKC_PHOSPHO_SITE 14-16; MYRISTYL 354-359; PKC_PHOSPHO_SITE 169-171; MYRISTYL 55-60; CAMP_PHOSPHO_SITE 193-196; CK2_PHOSPHO_SITE 625-628; MYRISTYL 367-372; ASN GLYCOSYLATION 578-581; MYRISTYL 61-66; PKC_PHOSPHO_SITE 464-466; CK2_PHOSPHO_SITE 531-534; MYRISTYL 228-233; MYRISTYL 579-584; TYR_PHOSPHO_SITE 581-587; MYRISTYL 214-219; CAMP_PHOSPHO_SITE 598-601; AMIDATION 191-194; ASN GLYCOSYLATION 12-15; CK2_PHOSPHO_SITE 396-399; MYRISTYL 9-14; AMIDATION 517-520; TYR_PHOSPHO_SITE 609-617;	IG_LIKE 86-237; ARG_RICH 485-606; TNFR_NGFR_1 267-304; CYS_RICH 280-304; ADH_SHORT 296-324; IG 96-239;
DEX0448_026.aa.2	N	0 - 01-388;	54-80,1.129; 13-23,1.141;	CK2_PHOSPHO_SITE 242-245; TYR_PHOSPHO_SITE 319-325; CAMP PHOSPHO SITE 336-339;	ADH_SHORT 34-62; ARG_RICH 223-344;

147

			119-128, 1.102; 29-50, 1.227; 353-362, 1.092; 187-192, 1.054; 140-147, 1.095; 229-241, 1.078; 108-116, 1.142; 377-385, 1.109;	CK2 PHOSPHO SITE 174-177; TYR PHOSPHO SITE 347-355; PKC PHOSPHO SITE 139-141; PKC PHOSPHO SITE 162-164; MYRISTYL 92-97; CK2 PHOSPHO SITE 363-366; AMIDATION 255-258; PKC PHOSPHO SITE 205-207; MYRISTYL 220-225; CK2 PHOSPHO SITE 38-41; TYR PHOSPHO SITE 284-290; ASN GLYCOSYLATION 316-319; MYRISTYL 255-260; TYR PHOSPHO SITE 268-274; MYRISTYL 317-322; PKC PHOSPHO SITE 202-204; CK2 PHOSPHO SITE 134-137; MYRISTYL 105-110; CK2 PHOSPHO SITE 269-272;	
DEX0448_027.aa.1	N	0 - ol-83;	8-15, 1.122; 35-51, 1.056; 61-70, 1.178; 17-23, 1.066;	MYRISTYL 49-54; PKC PHOSPHO SITE 70-72; MYRISTYL 15-20; CK2 PHOSPHO SITE 19-22;	
DEX0448_027.orf.1	N	0 - ol-379;	36-77, 1.182; 79-92, 1.088; 338-375, 1.175; 270-332, 1.197; 224-230, 1.069; 135-152, 1.163; 160-166, 1.055; 108-126, 1.15; 4-27, 1.171; 184-197, 1.183;	PKC PHOSPHO SITE 67-69; CK2 PHOSPHO SITE 237-240; MYRISTYL 91-96; CAMP PHOSPHO SITE 156-159; PKC PHOSPHO SITE 98-100; CK2 PHOSPHO SITE 262-265; PKC PHOSPHO SITE 155-157; MYRISTYL 321-326; PKC PHOSPHO SITE 254-256; CK2 PHOSPHO SITE 206-209; CK2 PHOSPHO SITE 269-272; PKC PHOSPHO SITE 154-156; CK2 PHOSPHO SITE 197-200; PKC PHOSPHO SITE 216-218; PKC PHOSPHO SITE 19-21; AMIDATION 232-235; CAMP PHOSPHO SITE 234-237; CK2 PHOSPHO SITE 202-205; CK2 PHOSPHO SITE 104-107; CK2 PHOSPHO SITE 204-207; ASN GLYCOSYLATION 323-326;	HLH 1 59-74; ORANGE 107-151; HLH 25-80; HLH 2 18-75; HLH 20-75;
DEX0448_027.orf.2	N	0 - ol-379;	160-166, 1.055; 108-126, 1.15; 184-197, 1.183; 79-92, 1.088; 135-152, 1.163; 224-230, 1.069;	MYRISTYL 91-96; PKC PHOSPHO SITE 155-157; ASN GLYCOSYLATION 323-326; PKC PHOSPHO SITE 19-21; CK2 PHOSPHO SITE 206-209; AMIDATION 232-235; CK2 PHOSPHO SITE 202-205; CAMP PHOSPHO SITE 156-159;	HLH 25-80; ORANGE 107-151; HLH 1 59-74; HLH 2 18-75; HLH 20-75;

			270-332, 1.197; 4-27, 1.171; 36- 77, 1.182; 338- 375, 1.175;	PKC_PHOSPHO_SITE 154-156; CK2_PHOSPHO_SITE 237-240; CK2_PHOSPHO_SITE 204-207; CK2_PHOSPHO_SITE 197-200; CK2_PHOSPHO_SITE 269-272; PKC_PHOSPHO_SITE 216-218; PKC_PHOSPHO_SITE 98-100; CAMP_PHOSPHO_SITE 234-237; CK2_PHOSPHO_SITE 262-265; MYRISTYL 321-326; PKC_PHOSPHO_SITE 254- 256; CK2_PHOSPHO_SITE 104-107; PKC_PHOSPHO_SITE 67-69;	
DEX0448_027.aa.3	N	0 - 01-358;	60-69, 1.117; 349-355, 1.057; 245-265, 1.091; 148-157, 1.12; 279-288, 1.174; 328-334, 1.079; 17-23, 1.066; 80-90, 1.123; 35-49, 1.127; 317-326, 1.107; 167-223, 1.149; 134-140, 1.039; 225-233, 1.102; 8-15, 1.122; 294-312, 1.179;	MYRISTYL 118-123; PKC_PHOSPHO_SITE 343- 345; CK2_PHOSPHO_SITE 262-265; PKC_PHOSPHO_SITE 121-123; CK2_PHOSPHO_SITE 268-271; PKC_PHOSPHO_SITE 122-124; MYRISTYL 216-221; PKC_PHOSPHO_SITE 233- 235; MYRISTYL 351-356; CK2_PHOSPHO_SITE 19-22; MYRISTYL 236-241; MYRISTYL 59-64; CK2_PHOSPHO_SITE 107-110; CAMP_PHOSPHO_SITE 53-56; PKC_PHOSPHO_SITE 249-251; CK2_PHOSPHO_SITE 95-98; CK2_PHOSPHO_SITE 286-289; PKC_PHOSPHO_SITE 219-221; PKC_PHOSPHO_SITE 332-334; CAMP_PHOSPHO_SITE 334-337; CK2_PHOSPHO_SITE 69-72; PKC_PHOSPHO_SITE 344-346; CK2_PHOSPHO_SITE 249-252; MYRISTYL 15-20;	
DEX0448_027.orf.3	N	0 - 01-329;	216-236, 1.091; 31-40, 1.117; 320-326, 1.057; 288-297, 1.107; 51-61, 1.123; 265-283, 1.179; 196-204, 1.102; 105-111, 1.039; 119-128, 1.12; 250-259, 1.174; 4-20, 1.063;	MYRISTYL 30-35; CK2_PHOSPHO_SITE 220-223; CK2_PHOSPHO_SITE 40-43; CAMP_PHOSPHO_SITE 24-27; CK2_PHOSPHO_SITE 257-260; MYRISTYL 187-192; PKC_PHOSPHO_SITE 92-94; CK2_PHOSPHO_SITE 66-69; MYRISTYL 207-212; CK2_PHOSPHO_SITE 233-236; PKC_PHOSPHO_SITE 4-6; CK2_PHOSPHO_SITE 78-81; PKC_PHOSPHO_SITE 315-317; PKC_PHOSPHO_SITE 190-192; CAMP_PHOSPHO_SITE 305-308; PKC_PHOSPHO_SITE 93-95; PKC_PHOSPHO_SITE 220-222; PKC_PHOSPHO_SITE 303-305;	

DEX0448_027.orf.4	N	0 - 01-425;	<p>299-305,1.079; 138-194,1.149;</p> <p>384-421,1.175; 125-138,1.088; 82-123,1.182; 181-198,1.163; 206-212,1.055; 65-73,1.09; 316-378,1.197; 44-52,1.102; 4- 10,1.067; 230- 243,1.183; 270- 276,1.069; 18- 33,1.14; 154- 172,1.15;</p>	<p>CK2_PHOSPHO_SITE 239-242; MYRISTYL 322-327; MYRISTYL 89-94; PKC_PHOSPHO_SITE 314-316; PKC_PHOSPHO_SITE 204-206;</p> <p>PKC_PHOSPHO_SITE 201-203; CK2_PHOSPHO_SITE 252-255; MYRISTYL 32-37; AMIDATION 278-281; CAMP_PHOSPHO_SITE 202-205; ASN_GLYCOSYLATION 369-372; CK2_PHOSPHO_SITE 248-251; CK2_PHOSPHO_SITE 308-311; MYRISTYL 55-60; PKC_PHOSPHO_SITE 200-202; CK2_PHOSPHO_SITE 283-286; CK2_PHOSPHO_SITE 250-253; PKC_PHOSPHO_SITE 144-146; PKC_PHOSPHO_SITE 113-115; CAMP_PHOSPHO_SITE 280-283; CK2_PHOSPHO_SITE 243-246; PKC_PHOSPHO_SITE 65-67; PKC_PHOSPHO_SITE 262-264; CK2_PHOSPHO_SITE 315-318; MYRISTYL 137-142; PKC_PHOSPHO_SITE 300-302; PKC_PHOSPHO_SITE 52-54; MYRISTYL 40-45; MYRISTYL 367-372; CK2_PHOSPHO_SITE 150-153;</p>	<p>HLH_1 105-120; HLH_2 57-121; HLH 66-121; HLH 71-126; ORANGE 153-197;</p>
DEX0448_027.orf.5	N	0 - 01-425;	<p>18-33,1.14; 125-138,1.088; 44-52,1.102; 154-172,1.15; 316-378,1.197; 65-73,1.09; 384-421,1.175; 270-276,1.069; 82-123,1.182; 4-10,1.067; 230-243,1.183; 206-212,1.055; 181-198,1.163;</p>	<p>CK2_PHOSPHO_SITE 252-255; CAMP_PHOSPHO_SITE 202-205; PKC_PHOSPHO_SITE 201-203; MYRISTYL 367-372; ASN_GLYCOSYLATION 369-372; CK2_PHOSPHO_SITE 250-253; PKC_PHOSPHO_SITE 262-264; MYRISTYL 40-45; CK2_PHOSPHO_SITE 315-318; PKC_PHOSPHO_SITE 300-302; CK2_PHOSPHO_SITE 308-311; CK2_PHOSPHO_SITE 243-246; CAMP_PHOSPHO_SITE 280-283; PKC_PHOSPHO_SITE 113-115; PKC_PHOSPHO_SITE 144-146; MYRISTYL 32-37; AMIDATION 278-281; PKC_PHOSPHO_SITE 65-67; CK2_PHOSPHO_SITE 248-251; PKC_PHOSPHO_SITE 52-54; PKC_PHOSPHO_SITE 200-202; MYRISTYL 55-60; MYRISTYL 137-142; CK2_PHOSPHO_SITE 150-153; CK2_PHOSPHO_SITE 283-286;</p>	<p>HLH 66-121; HLH 71-126; HLH_2 57-121; HLH_1 105-120; ORANGE 153-197;</p>

DEX0448_027.aa.6	N	0 - 01-64;	43-57,1.159; 4-9,1.165; 20-30,1.111;	ASN_GLYCOSYLATION 58-61; MYRISTYL 18-23; MYRISTYL 22-27; GLYCOSAMINOGLYCAN 15-18;	
DEX0448_027.orf.6	N	0 - 01-111;	4-9,1.072; 12-64,1.197; 70-107,1.175;	MYRISTYL 53-58; ASN_GLYCOSYLATION 55-58;	
DEX0448_028.aa.1	N	1 - 01-115;tm116-138;i139-145;	5-10,1.073; 81-91,1.159; 27-55,1.206; 108-142,1.208; 12-21,1.125; 95-106,1.112; 57-77,1.129;	MYRISTYL 127-132; MYRISTYL 118-123; PKC_PHOSPHO_SITE 17-19; CK2_PHOSPHO_SITE 61-64;	IG_MHC 84-90; IGc1 26-96; ig 24-88; IG_LIKE 7-99;
DEX0448_028.orf.1	Y	1 - 01-240;tm241-263;i264-270;	126-135,1.074; 74-81,1.071; 182-202,1.129; 233-267,1.208; 88-94,1.082; 220-231,1.112; 137-146,1.125; 206-216,1.159; 104-112,1.079; 152-180,1.206; 17-39,1.172; 50-71,1.187;	MYRISTYL 252-257; ASN_GLYCOSYLATION 50-53; MYRISTYL 243-248; MYRISTYL 67-72; PKC_PHOSPHO_SITE 142-144; CK2_PHOSPHO_SITE 186-189;	ig 149-213; IGc1 151-221; IG_MHC 209-215; MHC_II_alpha 50-133; IG_LIKE 130-224;
DEX0448_028.orf.2	Y	1 - 01-150;tm151-173;i174-180;	143-177,1.208; 130-141,1.112; 62-90,1.206; 36-45,1.074; 116-126,1.159; 4-22,1.247; 92-112,1.129; 47-56,1.125;	PKC_PHOSPHO_SITE 52-54; MYRISTYL 3-8; MYRISTYL 162-167; CK2_PHOSPHO_SITE 96-99; MYRISTYL 153-158;	ig 59-123; IG_MHC 119-125; IG_LIKE 40-134; IGc1 61-131;
DEX0448_029.aa.1	N	0 - 01-796;	508-553,1.18; 643-661,1.105; 597-603,1.132;	PKC_PHOSPHO_SITE 502-504; MYRISTYL 619-624; PKC_PHOSPHO_SITE 586-588; MYRISTYL 453-458; MYRISTYL 625-630; MYRISTYL 762-	GPROTEINBRPT 325-339; WD40 754-794; WD40 755-794; WD40

		<p>764-775, 1.122; 664-670, 1.113; 472-481, 1.114; 157-166, 1.115; 197-210, 1.113; 96-102, 1.04; 570-585, 1.178; 441-458, 1.189; 723-729, 1.052; 701-713, 1.181; 333-338, 1.052; 788-793, 1.169; 555-562, 1.057; 732-738, 1.063; 74-84, 1.198; 410-416, 1.069; 343-380, 1.196; 483-493, 1.082; 587-594, 1.062; 399-408, 1.24; 216-242, 1.234; 16-50, 1.144; 183-190, 1.101; 618-637, 1.233; 675-688, 1.078; 609-616, 1.15; 463-470, 1.095; 433-439, 1.108; 321-327, 1.154; 120-132, 1.088; 254-274, 1.152; 741-762, 1.2; 496-501, 1.078; 281-315, 1.231;</p>	<p>767; PKC_PHOSPHO_SITE 62-64; MYRISTYL 469-474; MYRISTYL 583-588; ASN GLYCOSYLATION 6-9; CK2_PHOSPHO_SITE 275-278; PKC_PHOSPHO_SITE 108-110; CK2_PHOSPHO_SITE 335-338; MYRISTYL 577-582; MYRISTYL 383-388; MYRISTYL 438-443; AMIDATION 639-642; CK2_PHOSPHO_SITE 406-409; MYRISTYL 320-325; MYRISTYL 270-275; ASN GLYCOSYLATION 147-150; CK2_PHOSPHO_SITE 193-196; CK2_PHOSPHO_SITE 2-5; PKC_PHOSPHO_SITE 275-277; PKC_PHOSPHO_SITE 684-686; CK2_PHOSPHO_SITE 11-14; PKC_PHOSPHO_SITE 387-389; TYR_PHOSPHO_SITE 201-208; CK2_PHOSPHO_SITE 20-23; MYRISTYL 61-66; PKC_PHOSPHO_SITE 758-760; MYRISTYL 100-105; MYRISTYL 442-447; PKC_PHOSPHO_SITE 199-201; MYRISTYL 501-506; PKC_PHOSPHO_SITE 167-169; LEUCINE_ZIPPER 26-47; CK2_PHOSPHO_SITE 239-242; MYRISTYL 510-515; CK2_PHOSPHO_SITE 162-165; CAMP_PHOSPHO_SITE 641-644; PKC_PHOSPHO_SITE 131-133; CK2_PHOSPHO_SITE 489-492; PKC_PHOSPHO_SITE 119-121; CK2_PHOSPHO_SITE 657-660; ASN GLYCOSYLATION 91-94; MYRISTYL 784-789; MYRISTYL 72-77; PKC_PHOSPHO_SITE 115-117; MYRISTYL 720-725; PKC_PHOSPHO_SITE 149-151; MYRISTYL 717-722; MYRISTYL 429-434; CK2_PHOSPHO_SITE 737-740; CK2_PHOSPHO_SITE 689-692;</p>	<p>242-280; WD40 642-681; WD_REPEATS_2_2 520-561; EF_HAND 327-339; WD_REPEATS_2_3 603-644; WD_REPEATS_2_4 649-690; HELP 164-241; GPROTEINBRPT 781-795; GPROTEINBRPT 539-553; WD40 341-380; WD_REPEATS_2_5 761-796; WD40 514-552; WD40 241-290; WD40 597-635; WD40 555-593; WD40 513-552; WD40 431-469; WD40 387-426; WD40 293-338; WD40 710-748; WD40 386-426; WD40 342-380; WD40 643-681; WD_REPEATS_REGION 329-796; WD40 430-469; WD40 596-635; WD_REPEATS_2_1 437-468; WD40 707-748;</p>
DEX0448_029.aa.2	N	<p>85-105, 1.152; 164-169, 1.052; 152-158, 1.154;</p>	<p>PKC_PHOSPHO_SITE 417-419; CK2_PHOSPHO_SITE 166-169; PKC_PHOSPHO_SITE 515-517; MYRISTYL 101-106; PKC_PHOSPHO_SITE 589-</p>	<p>WD_REPEATS_2_5 592-627; WD40 73-111; WD_REPEATS_2_2 351-</p>

		<p>554-560, 1.052; 386-393, 1.057; 449-468, 1.233; 532-544, 1.181; 572-593, 1.2; 241-247, 1.069; 230-239, 1.24; 47-73, 1.234; 506-519, 1.078; 440-447, 1.15; 418-425, 1.062; 272-289, 1.189; 112-146, 1.231; 339-384, 1.18; 563-569, 1.063; 28-41, 1.113; 595-606, 1.122; 174-211, 1.196; 428-434, 1.132; 401-416, 1.178; 303-312, 1.114; 327-332, 1.078; 474-492, 1.105; 619-624, 1.169; 264-270, 1.108; 314-324, 1.082; 294-301, 1.095; 495-501, 1.113; 14-21, 1.101;</p>	<p>591; MYRISTYL 414-419; MYRISTYL 332-337; MYRISTYL 456-461; MYRISTYL 548-553; MYRISTYL 300-305; PKC_PHOSPHO_SITE 333- 335; CAMP_PHOSPHO_SITE 472-475; CK2_PHOSPHO_SITE 237-240; MYRISTYL 341- 346; MYRISTYL 450-455; PKC_PHOSPHO_SITE 218-220; MYRISTYL 408-413; CK2_PHOSPHO_SITE 568-571; CK2_PHOSPHO_SITE 520-523; PKC_PHOSPHO_SITE 106-108; MYRISTYL 284-289; MYRISTYL 273-278; MYRISTYL 269-274; AMIDATION 470-473; TYR_PHOSPHO_SITE 32-39; MYRISTYL 615-620; CK2_PHOSPHO_SITE 24-27; MYRISTYL 214-219; CK2_PHOSPHO_SITE 488-491; MYRISTYL 260- 265; CK2_PHOSPHO_SITE 70-73; MYRISTYL 151- 156; MYRISTYL 551-556; CK2_PHOSPHO_SITE 106-109; CK2_PHOSPHO_SITE 320-323; PKC_PHOSPHO_SITE 30-32; MYRISTYL 593-598;</p>	<p>392; WD40 586-625; HELP 1-72; WD40 72- 121; GPROTEINRPT 370-384; GPROTEINRPT 156- 170; WD40 541-579; WD40 173-211; WD40 345-383; WD40 262- 300; WD40 428-466; GPROTEINRPT 612- 626; WD40 172-211; WD40 124-169; WD40 218-257; WD40 217- 257; EF HAND 158- 170; WD40 261-300; WD40 474-512; WD40 427-466; WD40 473- 512; WD_REPEATS_2_3 434-475; WD_REPEATS_2_4 480- 521; WD40 538-579; WD_REPEATS_REGION 160-627; WD40 386- 424; WD40 344-383; WD_REPEATS_2_1 268- 299; WD40 585-625;</p>
DEX0448_029.aa.3	N	<p>495-501, 1.113; 532-544, 1.181; 85-105, 1.152; 449-468, 1.233; 174-211, 1.196; 112-146, 1.231; 272-289, 1.189; 401-416, 1.178;</p>	<p>MYRISTYL 151-156; MYRISTYL 551-556; MYRISTYL 273-278; MYRISTYL 300-305; MYRISTYL 269-274; MYRISTYL 214-219; CAMP_PHOSPHO_SITE 472-475; CK2_PHOSPHO_SITE 24-27; MYRISTYL 456-461; MYRISTYL 260-265; CK2_PHOSPHO_SITE 70-73; CK2_PHOSPHO_SITE 237-240; MYRISTYL 593- 598; MYRISTYL 615-620; CK2 PHOSPHO SITE</p>	<p>WD40 124-169; WD40 586-625; WD40 585- 625; WD40 218-257; WD40 386-424; WD_REPEATS_2_3 434- 475; WD40 72-121; WD40 217-257; WD40 172-211;</p>

153

		<p>14-21, 1.101; 595-606, 1.122; 339-384, 1.18; 474-492, 1.105; 264-270, 1.108; 303-312, 1.114; 327-332, 1.078; 428-434, 1.132; 506-519, 1.078; 241-247, 1.069; 554-560, 1.052; 164-169, 1.052; 294-301, 1.095; 314-324, 1.082; 28-41, 1.113; 619-624, 1.169; 563-569, 1.063; 440-447, 1.15; 572-593, 1.2; 152-158, 1.154; 47-73, 1.234; 230-239, 1.24; 386-393, 1.057; 418-425, 1.062;</p>	<p>320-323; PKC_PHOSPHO_SITE 106-108; CK2_PHOSPHO_SITE 488-491; AMIDATION 470-473; PKC_PHOSPHO_SITE 333-335; MYRISTYL 548-553; CK2_PHOSPHO_SITE 166-169; PKC_PHOSPHO_SITE 417-419; CK2_PHOSPHO_SITE 106-109; PKC_PHOSPHO_SITE 218-220; MYRISTYL 341-346; MYRISTYL 408-413; PKC_PHOSPHO_SITE 589-591; PKC_PHOSPHO_SITE 30-32; PKC_PHOSPHO_SITE 515-517; MYRISTYL 332-337; CK2_PHOSPHO_SITE 568-571; MYRISTYL 284-289; MYRISTYL 414-419; TYR_PHOSPHO_SITE 32-39; MYRISTYL 101-106; CK2_PHOSPHO_SITE 520-523; MYRISTYL 450-455;</p>	<p>WD_REPEATS_2_2 351-392; WD_REPEATS_2_1 268-299; WD40 345-383; WD_REPEATS_2_4 480-521; WD40 428-466; EF_HAND 158-170; WD40 262-300; HELP 1-72; WD_REPEATS_2_5 592-627; GPROTEINBRPT 370-384; WD40 474-512; WD_REPEATS_REGION 160-627; GPROTEINBRPT 156-170; GPROTEINBRPT 612-626; WD40 538-579; WD40 173-211; WD40 261-300; WD40 73-111; WD40 427-466; WD40 344-383; WD40 473-512; WD40 541-579;</p>
DEX0448_029.aa.4	N	<p>237-244, 1.101; 453-462, 1.24; 150-156, 1.04; 464-470, 1.069; 174-186, 1.088; 308-328, 1.152; 397-434, 1.196; 475-480, 1.052; 211-220, 1.115; 5-10, 1.059; 251-264, 1.113; 128-138, 1.198; 41-56, 1.139;</p>	<p>PKC_PHOSPHO_SITE 329-331; MYRISTYL 29-34; PKC_PHOSPHO_SITE 441-443; PKC_PHOSPHO_SITE 169-171; MYRISTYL 154-159; ASN_GLYCOSYLATION 145-148; PKC_PHOSPHO_SITE 221-223; LEUCINE_ZIPPER 80-101; ASN_GLYCOSYLATION 201-204; ASN_GLYCOSYLATION 60-63; CK2_PHOSPHO_SITE 389-392; MYRISTYL 30-35; MYRISTYL 31-36; MYRISTYL 33-38; PKC_PHOSPHO_SITE 203-205; CK2_PHOSPHO_SITE 65-68; MYRISTYL 115-120; CK2_PHOSPHO_SITE 460-463; CK2_PHOSPHO_SITE 247-250; PKC_PHOSPHO_SITE 173-175; MYRISTYL 26-31; MYRISTYL 324-329;</p>	<p>WD40 396-434; WD40 296-334; HELP 218-295; WD40 295-344; WD40 347-392; WD40 440-480; WD40 395-434; EF_HAND 381-393; WD40 441-480;</p>

154

			375-381,1.154; 70-104,1.144; 387-392,1.052; 335-369,1.231; 270-296,1.234;	PKC_PHOSPHO_SITE 185-187; PKC_PHOSPHO_SITE 253-255; CK2_PHOSPHO_SITE 293-296; MYRISTYL 126-131; MYRISTYL 27-32; MYRISTYL 374-379; CK2_PHOSPHO_SITE 329-332; CK2_PHOSPHO_SITE 74-77; PKC_PHOSPHO_SITE 116-118; PKC_PHOSPHO_SITE 162-164; CK2_PHOSPHO_SITE 216-219; MYRISTYL 24-29; MYRISTYL 437-442; MYRISTYL 45-50; TYR_PHOSPHO_SITE 255-262;	
DEX0448_029.aa.5	N	1 - il- 70;tm71- 93;094-605;	379-394,1.178; 134-168,1.231; 361-376,1.127; 252-261,1.24; 573-584,1.122; 10-19,1.115; 107-127,1.152; 427-446,1.233; 406-412,1.132; 294-311,1.189; 349-354,1.078; 196-233,1.196; 186-191,1.052; 69-95,1.234; 510-522,1.181; 36-43,1.101; 452-470,1.105; 597-602,1.169; 316-323,1.095; 484-497,1.078; 418-425,1.15; 174-180,1.154; 473-479,1.113; 532-538,1.052; 263-269,1.069; 541-547,1.063; 286-292,1.108; 396-403,1.062;	MYRISTYL 363-368; MYRISTYL 282-287; PKC_PHOSPHO_SITE 493-495; CK2_PHOSPHO_SITE 188-191; CK2_PHOSPHO_SITE 46-49; MYRISTYL 291-296; CAMP_PHOSPHO_SITE 450-453; CK2_PHOSPHO_SITE 128-131; PKC_PHOSPHO_SITE 20-22; MYRISTYL 236-241; MYRISTYL 173-178; AMIDATION 448-451; MYRISTYL 386-391; MYRISTYL 392-397; CK2_PHOSPHO_SITE 546- 549; MYRISTYL 529-534; MYRISTYL 526-531; CK2_PHOSPHO_SITE 342-345; PKC_PHOSPHO_SITE 355-357; PKC_PHOSPHO_SITE 240-242; TYR_PHOSPHO_SITE 54-61; MYRISTYL 593-598; MYRISTYL 306-311; MYRISTYL 322-327; MYRISTYL 354-359; MYRISTYL 295-300; MYRISTYL 434-439; MYRISTYL 571-576; MYRISTYL 123-128; CK2_PHOSPHO_SITE 466- 469; CK2_PHOSPHO_SITE 15-18; PKC_PHOSPHO_SITE 567-569; MYRISTYL 428- 433; PKC_PHOSPHO_SITE 128-130; CK2_PHOSPHO_SITE 92-95; PKC_PHOSPHO_SITE 52-54; MYRISTYL 5-10; CK2_PHOSPHO_SITE 259-262; PKC_PHOSPHO_SITE 395-397; CK2_PHOSPHO_SITE 498-501;	WD_REPEATS_2_3 458- 499; WD40 519-557; WD40 451-490; WD40 195-233; WD40 95- 133; WD40 563-603; HELP 17-94; WD40 356-402; WD_REPEATS_2_2 412- 453; WD_REPEATS REGION 182-605; WD40 283- 322; GPROTEINBRPT 590-604; WD_REPEATS_2_1 290- 321; WD40 516-557; WD40 240-279; WD40 405-444; WD40 239- 279; GPROTEINBRPT 477-491; WD40 194- 233; GPROTEINBRPT 178-192; WD40 284- 322; EF_HAND 180- 192; WD40 94-143; WD40 146-191; WD40 564-603; WD40 452- 490; WD_REPEATS_2_4 570-605; WD40 406- 444;

155

				550-571,1.1.2; 325-334,1.114; 336-346,1.082; 50-63,1.113; 50-63,1.113; 408-415,1.057; 69-95,1.234; 450-456,1.132; 174-180,1.154; 462-482,1.15; 423-438,1.178; 484-492,1.185; 10-19,1.115; 196-233,1.196; 349-354,1.078; 252-261,1.24; 286-292,1.108; 361-406,1.18; 440-447,1.062; 263-269,1.069; 107-127,1.152; 336-346,1.082; 134-168,1.231; 186-191,1.052; 316-323,1.095; 294-311,1.189; 325-334,1.114; 36-43,1.101;			MYRISTYL 5-10; CK2_PHOSPHO_SITE 128-131; CK2 PHOSPHO SITE 15-18; MYRISTYL 430-435; MYRISTYL 295-300; CK2_PHOSPHO_SITE 92-95; MYRISTYL 322-327; PKC_PHOSPHO_SITE 439- 441; PKC_PHOSPHO_SITE 240-242; PKC_PHOSPHO_SITE 20-22; MYRISTYL 291-296; MYRISTYL 306-311; MYRISTYL 282-287; MYRISTYL 436-441; MYRISTYL 354-359; PKC_PHOSPHO_SITE 355-357; TYR_PHOSPHO_SITE 54-61; MYRISTYL 363-368; MYRISTYL 173-178; CK2_PHOSPHO_SITE 46-49; PKC PHOSPHO SITE 128-130; PKC_PHOSPHO_SITE 52-54; MYRISTYL 480-485; MYRISTYL 236-241; CK2 PHOSPHO SITE 342-345; CK2 PHOSPHO SITE 259-262; CK2 PHOSPHO_SITE 188-191; MYRISTYL 123-128;		WD40 194-233; WD_REPEATS_REGION_1 182-331; WD40 95- 133; WD40 195-233; GPROTEINBRPT 392- 406; WD40 367-405; WD40 240-279; GPROTEINBRPT 178- 192; WD40 408-446; WD_REPEATS_2_2 373- 414; EF HAND 180- 192; WD40 146-191; WD40 283-322; WD_REPEATS_2_1 290- 321; WD_REPEATS_REGION_2 373-414; GPROTEINBRPT 309- 323; HELP 17-94; WD40 284-322; WD40 366-405; WD40 94- 143; WD40 239-279;
DEX0448_029.aa.6	N	1 - il- 70;tm71- 93;094-495;		201-206,1.078; 295-302,1.057; 104-113,1.24; 327-334,1.062; 115-121,1.069; 415-428,1.078; 349-356,1.15; 472-478,1.063;		MYRISTYL 317-322; MYRISTYL 158-163; CK2 PHOSPHO SITE 477-480; PKC PHOSPHO_SITE 498-500; CK2 PHOSPHO_SITE 397-400; CK2_PHOSPHO_SITE 429-432; MYRISTYL 174- 179; MYRISTYL 7-12; MYRISTYL 323-328; PKC PHOSPHO_SITE 207-209; MYRISTYL 134- 139; MYRISTYL 206-211; PKC_PHOSPHO_SITE 92-94; CK2 PHOSPHO SITE 12-15; MYRISTYL	GPROTEINBRPT 521- 535; WD_REPEATS_2_2 260-301; WD_REPEATS_2_3 343- 384; WD40 253-292; WD_REPEATS_2_1 142- 173; WD_REPEATS_REGION 1		
DEX0448_029.orf.7	N	0 - ol-536;							

			<p>528-533, 1.169; 168-175, 1.095; 463-469, 1.052; 230-236, 1.031; 188-198, 1.082; 310-325, 1.178; 138-144, 1.108; 404-410, 1.113; 481-502, 1.2; 337-343, 1.132; 383-401, 1.105; 213-228, 1.127; 358-377, 1.233; 177-186, 1.114; 261-293, 1.129; 243-254, 1.141; 18-85, 1.231; 504-515, 1.122; 441-453, 1.181; 146-163, 1.189;</p>	<p>143-148; MYRISTYL 460-465; MYRISTYL 359-364; MYRISTYL 215-220; MYRISTYL 365-370; MYRISTYL 457-462; CK2 PHOSPHO SITE 194-197; AMIDATION 379-382; CK2 PHOSPHO SITE 111-114; MYRISTYL 88-93; PKC PHOSPHO SITE 326-328; CK2 PHOSPHO SITE 236-239; PKC PHOSPHO SITE 424-426; MYRISTYL 502-507; MYRISTYL 258-263; MYRISTYL 147-152; CAMP PHOSPHO SITE 381-384; MYRISTYL 524-529; PKC PHOSPHO SITE 12-14;</p>	<p>53-183; WD40 447-488; WD40 495-534; WD40 136-174; WD_REPEATS_2_5 501-536; WD_REPEATS_2_4 389-430; GPROTEINBRPT 279-293; WD40 254-292; WD40 47-85; GPROTEINBRPT 408-422; WD40 92-131; WD40 337-375; WD40 45-85; WD40 383-421; WD40 450-488; WD40 382-421; WD40 336-375; WD40 295-333; WD40 135-174; WD40 91-131; WD_REPEATS_REGION_2 260-536; WD40 494-534;</p>
<p>DEX0448_029.aa.7</p>	<p>N</p>	<p>0 - 01-544;</p>	<p>512-523, 1.122; 536-541, 1.169; 318-333, 1.178; 209-214, 1.078; 449-461, 1.181; 154-171, 1.189; 335-342, 1.062; 26-93, 1.231; 269-301, 1.129; 489-510, 1.2; 176-183, 1.095; 221-236, 1.127; 185-194, 1.114; 196-206, 1.082; 146-152, 1.108; 345-351, 1.132;</p>	<p>CK2 PHOSPHO SITE 119-122; MYRISTYL 214-219; MYRISTYL 182-187; MYRISTYL 151-156; MYRISTYL 155-160; PKC PHOSPHO SITE 215-217; MYRISTYL 15-20; CK2 PHOSPHO SITE 244-247; PKC PHOSPHO SITE 506-508; PKC PHOSPHO SITE 100-102; MYRISTYL 373-378; MYRISTYL 142-147; PKC PHOSPHO SITE 20-22; PKC PHOSPHO SITE 334-336; MYRISTYL 367-372; MYRISTYL 468-473; MYRISTYL 166-171; MYRISTYL 465-470; MYRISTYL 5-10; CK2 PHOSPHO SITE 202-205; CK2 PHOSPHO SITE 485-488; MYRISTYL 510-515; MYRISTYL 3-8; AMIDATION 387-390; MYRISTYL 532-537; CK2 PHOSPHO SITE 405-408; MYRISTYL 266-271; CAMP PHOSPHO SITE 389-392; PKC PHOSPHO SITE 432-434; MYRISTYL 96-101;</p>	<p>WD40 55-93; WD40 344-383; WD40 391-429; WD40 262-300; GPROTEINBRPT 416-430; GPROTEINBRPT 529-543; WD_REPEATS_2_2 268-309; WD40 100-139; WD40 458-496; WD40 503-542; WD_REPEATS_2_3 351-392; WD40 390-429; WD_REPEATS_2_1 150-181; WD40 502-542; WD40 143-182; GPROTEINBRPT 287-</p>

157

			123-129, 1.069; 423-436, 1.078; 366-385, 1.233; 303-310, 1.057; 251-262, 1.141; 480-486, 1.063; 112-121, 1.24; 412-418, 1.113; 471-477, 1.052; 357-364, 1.15; 238-244, 1.031; 391-409, 1.105;	CK2_PHOSPHO_SITE 20-23; MYRISTYL 325-330; CK2_PHOSPHO_SITE 437-440; MYRISTYL 223-228; MYRISTYL 331-336;	301; WD_REPEATS_2_5 509-544; WD40 99-139; WD_REPEATS_2_4 397-438; WD_REPEATS_REGION_1 61-191; WD40 53-93; WD40 144-182; WD40 455-496; WD40 345-383; WD40 303-341; WD_REPEATS_REGION_2 268-544; WD40 261-300;
DEX0448_030.orf.1	N	0 - 01-482;		CK2_PHOSPHO_SITE 62-65; CK2_PHOSPHO_SITE 383-386; CK2_PHOSPHO_SITE 450-453; LEUCINE_ZIPPER 388-409; PKC_PHOSPHO_SITE 184-186; CK2_PHOSPHO_SITE 124-127; MYRISTYL 287-292; PKC_PHOSPHO_SITE 137-139; CK2_PHOSPHO_SITE 204-207; PKC_PHOSPHO_SITE 269-271; CK2_PHOSPHO_SITE 432-435; MYRISTYL 299-304; CK2_PHOSPHO_SITE 56-59; ASN_GLYCOSYLATION 322-325; CK2_PHOSPHO_SITE 403-406; PKC_PHOSPHO_SITE 276-278; MYRISTYL 89-94; PKC_PHOSPHO_SITE 204-206; MYRISTYL 472-477; CAMP_PHOSPHO_SITE 263-266; CK2_PHOSPHO_SITE 358-361; CK2_PHOSPHO_SITE 233-236; PKC_PHOSPHO_SITE 337-339; ASN_GLYCOSYLATION 231-234; MYRISTYL 95-100; LEUCINE_ZIPPER 395-416; CK2_PHOSPHO_SITE 308-311; MYRISTYL 91-96;	
DEX0448_030.aa.1	N	0 - 01-454;	442-449, 1.086; 352-357, 1.069; 139-148, 1.118; 182-188, 1.068; 25-32, 1.077; 48-57, 1.129;	CK2_PHOSPHO_SITE 280-283; CK2_PHOSPHO_SITE 355-358; PKC_PHOSPHO_SITE 248-250; MYRISTYL 63-68; MYRISTYL 271-276; PKC_PHOSPHO_SITE 241-243; MYRISTYL 259-264; PKC_PHOSPHO_SITE 176-178; CK2_PHOSPHO_SITE 330-333; PKC_PHOSPHO_SITE	

			105-130,1.126; 335-346,1.105; 70-81,1.193; 425-432,1.074; 414-422,1.103; 259-265,1.056; 206-216,1.094; 363-387,1.133; 219-224,1.047; 236-247,1.095; 10-19,1.097; 168-173,1.066;	309-311; PKC_PHOSPHO_SITE 109-111; MYRISTYL 67-72; CAMP_PHOSPHO_SITE 235-238; CK2_PHOSPHO_SITE 34-37; ASN_GLYCOSYLATION 203-206; CK2_PHOSPHO_SITE 205-208; LEUCINE_ZIPPER 367-388; ASN_GLYCOSYLATION 294-297; MYRISTYL 61-66; CK2_PHOSPHO_SITE 96-99; MYRISTYL 444-449; CK2_PHOSPHO_SITE 404-407; CK2_PHOSPHO_SITE 422-425; LEUCINE_ZIPPER 360-381; CK2_PHOSPHO_SITE 375-378; CK2_PHOSPHO_SITE 176-179; PKC_PHOSPHO_SITE 156-158; CK2_PHOSPHO_SITE 28-31;		NDK 55-193; SP_Q9Y5B8_NDK7_HUMA N 56-187; NDK 201- 339; SP_Q9Y5B8_NDK7_HUMA N 207-336; NUCDPKINASE 100- 119; NDK 56-203; DM10 1-55; NUCDPKINASE 120- 137; NUCDPKINASE 167-186;
DEX0448_031.aa.1	N	0 - 01-340;	325-337,1.143; 268-275,1.139; 286-292,1.062; 58-67,1.088; 29-43,1.154; 130-137,1.069; 200-225,1.23; 111-122,1.068; 98-107,1.095; 315-323,1.141; 248-262,1.146; 184-192,1.056;	ASN_GLYCOSYLATION 202-205; TYR_PHOSPHO_SITE 242-249; MYRISTYL 146- 151; PKC_PHOSPHO_SITE 93-95; CK2_PHOSPHO_SITE 14-17; CK2_PHOSPHO_SITE 284-287; MYRISTYL 164-169; CK2_PHOSPHO_SITE 53-56; CK2_PHOSPHO_SITE 93-96; ASN_GLYCOSYLATION 280-283; MYRISTYL 191-196; PKC_PHOSPHO_SITE 53-55; PKC_PHOSPHO_SITE 155-157; PKC_PHOSPHO_SITE 197-199; MYRISTYL 27-32; MYRISTYL 303-308; PKC_PHOSPHO_SITE 284-286; PKC_PHOSPHO_SITE 47-49; PKC_PHOSPHO_SITE 304-306; MYRISTYL 136-141; MYRISTYL 285-290; PKC_PHOSPHO_SITE 229-231; PKC_PHOSPHO_SITE 248-250; CK2_PHOSPHO_SITE 256-259; PKC_PHOSPHO_SITE 148-150; MYRISTYL 115- 120; MYRISTYL 294-299; PKC_PHOSPHO_SITE 276-278; CAMP_PHOSPHO_SITE 75-78; MYRISTYL 12-17; AMIDATION 95-98; ASN_GLYCOSYLATION 271-274; MYRISTYL 60-65; PKC_PHOSPHO_SITE 159-161; AMIDATION 236-239; MYRISTYL 241- 246; PKC_PHOSPHO_SITE 280-282; MYRISTYL 172-177; CK2_PHOSPHO_SITE 111-114; MYRISTYL 264-269; MYRISTYL 86-91;		RIBOSOMAL_L2 202- 213; Ribosomal_L2 16-95; Ribosomal_L2_C 101- 236;
DEX0448_032.orf.1	N	0 - 01-304;				

DEX0448_032.aa.1	N	0 - 01-185;	90-96,1.073; 56-66,1.219; 41-54,1.075; 98-104,1.062; 4-12,1.199; 153-160,1.097; 126-134,1.102; 27-33,1.112; 71-83,1.059;	MYRISTYL 130-135; MYRISTYL 61-66; MYRISTYL 25-30; CK2_PHOSPHO_SITE 177-180; PKC_PHOSPHO_SITE 48-50; MYRISTYL 175-180; CK2_PHOSPHO_SITE 145-148; PKC_PHOSPHO_SITE 137-139; PKC_PHOSPHO_SITE 118-120; PKC_PHOSPHO_SITE 37-39; MYRISTYL 4-9; AMIDATION 125-128;	RIBOSOMAL_L2_91-102; RIBOSOMAL_L2_C1-125;
DEX0448_033.orf.1	N	0 - 01-484;	102-120,1.266; 323-341,1.148; 123-129,1.054; 464-469,1.084; 389-420,1.2; 245-253,1.153; 308-314,1.083; 260-267,1.107; 175-182,1.165; 206-212,1.073; 185-194,1.167; 225-239,1.095; 15-67,1.085; 422-448,1.166; 475-481,1.1; 92-98,1.085; 148-166,1.165; 75-84,1.175; 349-375,1.153; 378-386,1.148; 274-279,1.05;	PKC_PHOSPHO_SITE 463-465; AMIDATION 214-217; MYRISTYL 14-19; CK2_PHOSPHO_SITE 71-74; CK2_PHOSPHO_SITE 320-323; CK2_PHOSPHO_SITE 375-378; MYRISTYL 399-404; AMIDATION 1-4; CK2_PHOSPHO_SITE 321-324; PKC_PHOSPHO_SITE 84-87; MYRISTYL 89-94; CK2_PHOSPHO_SITE 244-246; CK2_PHOSPHO_SITE 189-192; MYRISTYL 339-344; CK2_PHOSPHO_SITE 201-204; MYRISTYL 94-99; ASN_GLYCOSYLATION 379-382; MYRISTYL 92-97; MYRISTYL 341-346; AMIDATION 7-10;	MR_MLE_2 288-319; MR_MLE_N 49-180; MR_MLE 217-468;
DEX0448_033.aa.1	N	0 - 01-483;	4-66,1.085; 307-313,1.083; 174-181,1.165; 224-238,1.095; 421-447,1.166; 91-97,1.085;	MYRISTYL 88-93; AMIDATION 213-216; PKC_PHOSPHO_SITE 369-371; CK2_PHOSPHO_SITE 83-86; CK2_PHOSPHO_SITE 319-322; MYRISTYL 340-345; CK2_PHOSPHO_SITE 320-323; MYRISTYL 338-343; CK2_PHOSPHO_SITE 374-377; PKC_PHOSPHO_SITE 462-464; MYRISTYL	MR_MLE_N 48-179; ALA_RICH 4-67; MR_MLE 216-467; MR_MLE_2 287-318;

160

			377-385,1.148; 147-165,1.165; 205-211,1.073; 259-266,1.107; 463-468,1.084; 184-193,1.167; 474-480,1.1; 101-119,1.266; 273-278,1.05; 322-340,1.148; 244-252,1.153; 122-128,1.054; 388-419,1.2; 74-83,1.175; 348-374,1.153; 422-448,1.166; 225-239,1.095; 206-212,1.073; 102-120,1.266; 323-341,1.148; 185-194,1.167; 378-386,1.148; 15-67,1.085; 274-279,1.05; 175-182,1.165; 260-267,1.107; 464-469,1.084; 123-129,1.054; 75-84,1.175; 349-375,1.153; 148-166,1.165; 245-253,1.153; 92-98,1.085; 308-314,1.083; 389-420,1.2; 475-481,1.1;	93-98; MYRISTYL 398-403; CK2_PHOSPHO_SITE 70-73; PKC_PHOSPHO_SITE 243-245; MYRISTYL 91-96; ASN_GLYCOSYLATION 378-381; CK2_PHOSPHO_SITE 200-203; CK2_PHOSPHO_SITE 188-191;	
DEX0448_033.orf.2	N	0 - 01-484;	PKC_PHOSPHO_SITE 370-372; CK2_PHOSPHO_SITE 71-74; CK2_PHOSPHO_SITE 84-87; AMIDATION 7-10; MYRISTYL 92-97; AMIDATION 214-217; MYRISTYL 94-99; CK2_PHOSPHO_SITE 201-204; MYRISTYL 399-404; ASN_GLYCOSYLATION 379- 382; CK2_PHOSPHO_SITE 375-378; MYRISTYL 341-346; CK2_PHOSPHO_SITE 320-323; MYRISTYL 14-19; MYRISTYL 339-344; PKC_PHOSPHO_SITE 244-246; CK2_PHOSPHO_SITE 321-324; CK2_PHOSPHO_SITE 189-192; AMIDATION 1-4; MYRISTYL 89-94; PKC_PHOSPHO_SITE 463-465;	MR_MLE 217-468; MR_MLE_N 49-180; MR_MLE_2 288-319;	

DEX0448_033.orf.3	N	0 - 01-484;	<p>15-67,1.085; 389-420,1.2; 175-182,1.165; 323-341,1.148; 349-375,1.153; 123-129,1.054; 464-469,1.084; 75-84,1.175; 245-253,1.153; 92-98,1.085; 148-166,1.165; 378-386,1.148; 225-239,1.095; 308-314,1.083; 475-481,1.1; 274-279,1.05; 185-194,1.167; 102-120,1.266; 206-212,1.073; 260-267,1.107; 422-448,1.166;</p>	<p>PKC PHOSPHO_SITE 370-372; MYRISTYL 92-97; AMIDATION 214-217; CK2_PHOSPHO_SITE 320-323; PKC_PHOSPHO_SITE 463-465; MYRISTYL 14-19; CK2_PHOSPHO_SITE 375-378; CK2_PHOSPHO_SITE 71-74; CK2_PHOSPHO_SITE 201-204; MYRISTYL 339-344; PKC PHOSPHO_SITE 244-246; AMIDATION 7-10; CK2_PHOSPHO_SITE 84-87; MYRISTYL 341-346; AMIDATION 1-4; MYRISTYL 94-99; MYRISTYL 399-404; ASN GLYCOSYLATION 379-382; CK2_PHOSPHO_SITE 321-324; MYRISTYL 89-94; CK2_PHOSPHO_SITE 189-192;</p>	<p>MR_MLE_N 49-180; MR_MLE_2 288-319; MR_MLE 217-468;</p>
DEX0448_033.orf.4	N	0 - 01-249;	<p>39-44,1.05; 143-151,1.148; 154-185,1.2; 88-106,1.148; 24-32,1.107; 187-213,1.166; 229-234,1.084; 73-79,1.083; 114-140,1.153; 240-246,1.1;</p>	<p>CK2 PHOSPHO_SITE 85-88; PKC PHOSPHO_SITE 135-137; CK2_PHOSPHO_SITE 86-89; MYRISTYL 164-169; PKC_PHOSPHO_SITE 228-230; PKC PHOSPHO_SITE 16-18; MYRISTYL 104-109; MYRISTYL 106-111; CK2_PHOSPHO_SITE 140-143; ASN GLYCOSYLATION 144-147;</p>	<p>MR_MLE_2 53-84; MR_MLE 2-233;</p>
DEX0448_033.aa.4	N	0 - 01-221;	<p>91-97,1.085; 184-193,1.167; 122-128,1.054; 4-66,1.085; 174-181,1.165;</p>	<p>CK2_PHOSPHO_SITE 200-203; CK2_PHOSPHO_SITE 70-73; CK2_PHOSPHO_SITE 188-191; AMIDATION 213-216; CK2_PHOSPHO_SITE 83-86; MYRISTYL 93-98; MYRISTYL 91-96; MYRISTYL 88-93;</p>	<p>MR_MLE_N 48-179; ALA_RICH 4-67;</p>

DEX0448_033.orf.5	N	0 - 01-416;	<p>74-83,1.175; 205-211,1.073; 101-119,1.266; 147-165,1.165; 404-413,1.123; 102-120,1.266; 396-401,1.059; 225-239,1.095; 123-129,1.054; 323-341,1.148; 148-166,1.165; 15-67,1.085; 185-194,1.167; 274-279,1.05; 370-394,1.156; 175-182,1.165; 206-212,1.073; 75-84,1.175; 92-98,1.085; 245-253,1.153; 260-267,1.107; 308-314,1.083;</p>	<p>MYRISTYL 94-99; CK2_PHOSPHO_SITE 201-204; CK2_PHOSPHO_SITE 189-192; AMIDATION 7-10; MYRISTYL 89-94; CK2_PHOSPHO_SITE 321-324; CK2_PHOSPHO_SITE 71-74; CK2_PHOSPHO_SITE 402-405; MYRISTYL 339-344; PKC_PHOSPHO_SITE 367-369; AMIDATION 1-4; AMIDATION 214-217; MYRISTYL 14-19; MYRISTYL 376-381; CK2_PHOSPHO_SITE 84-87; CK2_PHOSPHO_SITE 320-323; MYRISTYL 92-97; MYRISTYL 341-346; PKC_PHOSPHO_SITE 244- 246;</p>	<p>MR_MLE_2 288-319; MR_MLE_N 49-180;</p>
DEX0448_033.aa.5	N	0 - 01-415;	<p>174-181,1.165; 307-313,1.083; 205-211,1.073; 369-393,1.156; 395-400,1.059; 91-97,1.085; 184-193,1.167; 122-128,1.054; 403-412,1.123; 322-340,1.148; 224-238,1.095; 259-266,1.107; 244-252,1.153; 74-83,1.175;</p>	<p>MYRISTYL 340-345; CK2_PHOSPHO_SITE 200- 203; MYRISTYL 91-96; CK2_PHOSPHO_SITE 83- 86; CK2_PHOSPHO_SITE 188-191; MYRISTYL 93- 98; CK2_PHOSPHO_SITE 70-73; MYRISTYL 375- 380; MYRISTYL 338-343; MYRISTYL 88-93; AMIDATION 213-216; CK2_PHOSPHO_SITE 401- 404; PKC_PHOSPHO_SITE 243-245; CK2_PHOSPHO_SITE 320-323; PKC_PHOSPHO_SITE 366-368; CK2_PHOSPHO_SITE 319-322;</p>	<p>ALA_RICH 4-67; MR_MLE_2 287-318; MR_MLE_N 48-179;</p>

163

DEX0448_033.orf.6	N	0 - 01-484;	<p>273-278,1.05; 4-66,1.085; 147-165,1.165; 101-119,1.266; 92-98,1.085; 378-386,1.148; 75-84,1.175; 349-375,1.153; 206-212,1.073; 422-448,1.166; 225-239,1.095; 274-279,1.05; 185-194,1.167; 123-129,1.054; 260-267,1.107; 475-481,1.1; 389-420,1.2; 15-67,1.085; 245-253,1.153; 148-166,1.165; 308-314,1.083; 102-120,1.266; 464-469,1.084; 323-341,1.148; 175-182,1.165;</p>	<p>MYRISTYL 14-19; MYRISTYL 399-404; ASN GLYCOSYLATION 379-382; PKC_PHOSPHO_SITE 244-246; CK2_PHOSPHO_SITE 375-378; CK2_PHOSPHO_SITE 189-192; AMIDATION 1-4; PKC_PHOSPHO_SITE 463-465; MYRISTYL 341-346; MYRISTYL 89-94; CK2_PHOSPHO_SITE 71-74; CK2_PHOSPHO_SITE 84-87; AMIDATION 214-217; MYRISTYL 94-99; CK2_PHOSPHO_SITE 321-324; MYRISTYL 92-97; PKC_PHOSPHO_SITE 370-372; CK2_PHOSPHO_SITE 320-323; MYRISTYL 339-344; AMIDATION 7-10; CK2_PHOSPHO_SITE 201-204;</p>	<p>MR_MLE_2 217-468; MR_MLE_2 288-319; MR_MLE_N 49-180;</p>
DEX0448_033.orf.7	N	0 - 01-349;	<p>75-84,1.175; 92-98,1.085; 185-194,1.167; 15-67,1.085; 102-120,1.266; 225-239,1.095; 308-314,1.083; 274-279,1.05; 323-341,1.148; 148-166,1.165; 206-212,1.073;</p>	<p>MYRISTYL 92-97; CK2_PHOSPHO_SITE 71-74; PKC_PHOSPHO_SITE 244-246; CK2_PHOSPHO_SITE 321-324; MYRISTYL 339-344; CK2_PHOSPHO_SITE 201-204; CK2_PHOSPHO_SITE 189-192; MYRISTYL 341-346; CK2_PHOSPHO_SITE 320-323; AMIDATION 214-217; MYRISTYL 94-99; AMIDATION 7-10; MYRISTYL 14-19; AMIDATION 1-4; MYRISTYL 89-94; CK2_PHOSPHO_SITE 84-87;</p>	<p>MR_MLE_2 288-319; MR_MLE_N 49-180;</p>

				245-253,1.153; 123-129,1.054; 175-182,1.165; 260-267,1.107;			
DEX0448_033.aa.7	N		0 - 01-305;	79-85,1.054; 162-168,1.073; 104-122,1.165; 201-209,1.153; 4-23,1.085; 131-138,1.165; 48-54,1.085; 264-270,1.083; 141-150,1.167; 181-195,1.095; 279-297,1.148; 230-235,1.05; 31-40,1.175; 58-76,1.266; 216-223,1.107;	CK2_PHOSPHO_SITE 145-148; AMIDATION 170-173; CK2_PHOSPHO_SITE 157-160; MYRISTYL 295-300; CK2_PHOSPHO_SITE 40-43; CK2_PHOSPHO_SITE 277-280; MYRISTYL 50-55; MYRISTYL 45-50; MYRISTYL 48-53; MYRISTYL 297-302; CK2_PHOSPHO_SITE 27-30; CK2_PHOSPHO_SITE 276-279; PKC_PHOSPHO_SITE 200-202;	MR_MLE_N 5-136; MR_MLE_2 244-275;	
DEX0448_033.orf.8	N		0 - 01-484;	185-194,1.167; 175-182,1.165; 475-481,1.1; 260-267,1.107; 422-448,1.166; 123-129,1.054; 75-84,1.175; 274-279,1.05; 349-375,1.153; 225-239,1.095; 206-212,1.073; 102-120,1.266; 245-253,1.153; 92-98,1.085; 378-386,1.148; 389-420,1.2; 464-469,1.084;	CK2_PHOSPHO_SITE 321-324; MYRISTYL 94-99; MYRISTYL 339-344; PKC_PHOSPHO_SITE 370-372; CK2_PHOSPHO_SITE 320-323; MYRISTYL 14-19; CK2_PHOSPHO_SITE 201-204; MYRISTYL 89-94; AMIDATION 7-10; CK2_PHOSPHO_SITE 84-87; AMIDATION 1-4; PKC_PHOSPHO_SITE 244-246; MYRISTYL 399-404; CK2_PHOSPHO_SITE 375-378; CK2_PHOSPHO_SITE 71-74; MYRISTYL 341-346; PKC_PHOSPHO_SITE 463-465; AMIDATION 214-217; ASN GLYCOSYLATION 379-382; CK2_PHOSPHO_SITE 189-192; MYRISTYL 92-97;	MR_MLE 217-468; MR_MLE_N 49-180; MR_MLE_2 288-319;	

165

			148-166,1.165; 308-314,1.083; 323-341,1.148; 15-67,1.085;			
DEX0448_033.orf.9	N	0 - 01-484;	260-267,1.107; 349-375,1.153; 15-67,1.085; 185-194,1.167; 225-239,1.095; 175-182,1.165; 274-279,1.05; 475-481,1.1; 102-120,1.266; 378-386,1.148; 92-98,1.085; 206-212,1.073; 148-166,1.165; 323-341,1.148; 123-129,1.054; 75-84,1.175; 422-448,1.166; 245-253,1.153; 464-469,1.084; 389-420,1.2; 308-314,1.083;	MYRISTYL 94-99; CK2_PHOSPHO_SITE 375-378; MYRISTYL 339-344; PKC_PHOSPHO_SITE 244-246; MYRISTYL 341-346; MYRISTYL 89-94; ASN GLYCOSYLATION 379-382; CK2_PHOSPHO_SITE 189-192; MYRISTYL 14-19; PKC_PHOSPHO_SITE 370-372; CK2_PHOSPHO_SITE 320-323; AMIDATION 1-4; MYRISTYL 92-97; AMIDATION 7-10; PKC_PHOSPHO_SITE 463-465; CK2_PHOSPHO_SITE 84-87; CK2_PHOSPHO_SITE 201-204; AMIDATION 214-217; MYRISTYL 399-404; CK2_PHOSPHO_SITE 71-74; CK2_PHOSPHO_SITE 321-324;	MR_MLE_N 49-180; MR_MLE 217-468; MR_MLE_2 288-319;	
DEX0448_033.orf.10	N	0 - 01-484;	422-448,1.166; 349-375,1.153; 308-314,1.083; 15-67,1.085; 323-341,1.148; 75-84,1.175; 92-98,1.085; 378-386,1.148; 389-420,1.2; 464-469,1.084; 245-253,1.153;	MYRISTYL 92-97; AMIDATION 7-10; MYRISTYL 341-346; PKC_PHOSPHO_SITE 463-465; CK2_PHOSPHO_SITE 71-74; MYRISTYL 14-19; CK2_PHOSPHO_SITE 84-87; CK2_PHOSPHO_SITE 375-378; PKC_PHOSPHO_SITE 370-372; CK2_PHOSPHO_SITE 201-204; AMIDATION 214-217; MYRISTYL 89-94; CK2_PHOSPHO_SITE 320-323; MYRISTYL 339-344; PKC_PHOSPHO_SITE 244-246; AMIDATION 1-4; CK2_PHOSPHO_SITE 189-192; CK2_PHOSPHO_SITE 321-324; ASN GLYCOSYLATION 379-382; MYRISTYL 399-	MR_MLE 217-468; MR_MLE_2 288-319; MR_MLE_N 49-180;	

			206-212, 1.073; 225-239, 1.095; 102-120, 1.266; 185-194, 1.167; 175-182, 1.165; 260-267, 1.107; 274-279, 1.05; 148-166, 1.165; 475-481, 1.1; 123-129, 1.054;	404; MYRISTYL 94-99;		
DEX0448_034.aa.1	Y	0 - 01-138;	34-47, 1.124; 54-62, 1.107; 117-125, 1.1; 88-102, 1.144; 109-115, 1.085; 4-26, 1.154; 69- 76, 1.073;	MYRISTYL 115-120; PKC_PHOSPHO_SITE 133-135; PKC_PHOSPHO_SITE 73-75; CK2_PHOSPHO_SITE 92-95; PKC_PHOSPHO_SITE 46-48; MYRISTYL 13-18; PKC_PHOSPHO_SITE 33-35;	Ribosomal_S17e 5-125; RIBOSOMAL_S17E 44-59;	
DEX0448_035.aa.1	N	0 - 01-386;	36-45, 1.073; 238-244, 1.063; 283-295, 1.148; 115-138, 1.126; 214-226, 1.133; 187-196, 1.129; 307-316, 1.067; 158-169, 1.16; 84-101, 1.104; 25-33, 1.089; 198-204, 1.049; 346-383, 1.206; 174-183, 1.094; 50-78, 1.102; 10-20, 1.2; 247-254, 1.034;	PKC_PHOSPHO_SITE 32-34; CK2_PHOSPHO_SITE 140-143; ASN_GLYCOSYLATION 322-325; CK2_PHOSPHO_SITE 281-284; PKC_PHOSPHO_SITE 112-114; MYRISTYL 183-188; CK2_PHOSPHO_SITE 336-339; CK2_PHOSPHO_SITE 64-67; MYRISTYL 51-56; MYRISTYL 195-200; MYRISTYL 79-84; MYRISTYL 320-325; PKC_PHOSPHO_SITE 281-283; ASN_GLYCOSYLATION 170-173; MYRISTYL 107-112; MYRISTYL 263-268; PKC_PHOSPHO_SITE 64-66; CK2_PHOSPHO_SITE 381-384; MYRISTYL 323-328;	MEVGALKINASE 320-337; MEVGALKINASE 174-196; PRICHEXTENSIN 119-136; PRICHEXTENSIN 83-104; PRICHEXTENSIN 151-176; PRICHEXTENSIN 49-61;	
DEX0448_035.aa.4	N	0 - 01-492;	36-45, 1.073; 238-244, 1.063; 115-138, 1.126;	CK2_PHOSPHO_SITE 374-377; PKC_PHOSPHO_SITE 64-66; CK2_PHOSPHO_SITE 140-143; CK2 PHOSPHO SITE 281-284; PKC PHOSPHO SITE	PRICHEXTENSIN 352-364; PRICHEXTENSIN 364-385;	

			<p>477-489, 1.11; 50-78, 1.102; 440-449, 1.068; 10-20, 1.2; 84- 101, 1.104; 158- 169, 1.16; 247- 254, 1.034; 198- 204, 1.049; 174- 183, 1.094; 214- 226, 1.133; 25- 33, 1.089; 429- 438, 1.123; 457- 474, 1.129; 283- 295, 1.148; 307- 314, 1.114; 356- 367, 1.106; 405- 422, 1.132; 386- 394, 1.175; 187- 196, 1.129;</p>	<p>32-34; CK2_PHOSPHO_SITE 343-346; MYRISTYL 398-403; CK2_PHOSPHO_SITE 488-491; CK2_PHOSPHO_SITE 64-67; MYRISTYL 51-56; MYRISTYL 397-402; PKC_PHOSPHO_SITE 438- 440; PKC_PHOSPHO_SITE 281-283; MYRISTYL 195-200; PKC_PHOSPHO_SITE 434-436; MYRISTYL 183-188; MYRISTYL 107-112; PKC_PHOSPHO_SITE 112-114; ASN_GLYCOSYLATION 371-374; MYRISTYL 79-84; PKC_PHOSPHO_SITE 370-372; ASN_GLYCOSYLATION 170-173; CAMP_PHOSPHO_SITE 318-321; PKC_PHOSPHO_SITE 314-316; MYRISTYL 263- 268;</p>	<p>PRICHEXTENSIN 314- 326;</p>
DEX0448_036.aa.1	N	0 - 01-358;	<p>337-355, 1.182; 34-46, 1.187; 80-88, 1.119; 233-239, 1.061; 52-69, 1.095; 120-129, 1.171; 186-199, 1.139; 145-158, 1.116; 166-183, 1.148; 134-139, 1.069; 259-268, 1.091; 13-32, 1.162; 286-311, 1.137; 275-283, 1.087; 101-115, 1.176; 201-229, 1.192; 248-253, 1.079;</p>	<p>CAMP_PHOSPHO_SITE 130-133; CK2_PHOSPHO_SITE 113-116; CK2_PHOSPHO_SITE 211-214; MYRISTYL 199-204; PKC_PHOSPHO_SITE 240-242; CK2_PHOSPHO_SITE 98-101; PKC_PHOSPHO_SITE 329-331; CK2_PHOSPHO_SITE 47-50; CK2_PHOSPHO_SITE 71-74; PKC_PHOSPHO_SITE 253-255; MYRISTYL 175-180; CK2_PHOSPHO_SITE 89-92; PKC_PHOSPHO_SITE 335-337; MYRISTYL 103- 108;</p>	<p>D_2_HYDROXYACID_DH_ 1 166-193; 2- Hacid_DH 18-111; 2- Hacid_DH_C 113-304;</p>

168

DEX0448_036.aa.2	N	0 - 01-364;	241-268, 1.191; 74-91, 1.116; 39-47, 1.064; 93-107, 1.148; 313-334, 1.258; 287-302, 1.122; 341-361, 1.292; 271-280, 1.125; 12-23, 1.136; 217-224, 1.157; 195-205, 1.103; 207-212, 1.058; 25-33, 1.074; 58-66, 1.119; 4- 10, 1.027; 137- 173, 1.158; 122- 134, 1.126;	MYRISTYL 354-359; MYRISTYL 227-232; CK2_PHOSPHO_SITE 10-13; MYRISTYL 317-322; MYRISTYL 339-344; MYRISTYL 6-11; PKC_PHOSPHO_SITE 191-193; PKC_PHOSPHO_SITE 281-283; PKC_PHOSPHO_SITE 197-199; MYRISTYL 233-238; PKC_PHOSPHO_SITE 234- 236; MYRISTYL 309-314;	
DEX0448_036.orf.2	N	0 - 01-192;	38-74, 1.158; 108-113, 1.058; 118-125, 1.157; 27-35, 1.076; 131-182, 1.151; 11-17, 1.091; 96-106, 1.103; 101-116, 1.122; 31-38, 1.157; 55-82, 1.191; 127-148, 1.258; 155-175, 1.292; 6-27, 1.137; 85- 94, 1.125;	MYRISTYL 12-17; MYRISTYL 23-28; MYRISTYL 151-156; PKC_PHOSPHO_SITE 92-94; MYRISTYL 139-144; PKC_PHOSPHO_SITE 98-100; PKC_PHOSPHO_SITE 178-180; MYRISTYL 144- 149;	PRICHEXTENS 165- 177; PRICHEXTENS 18-30;
DEX0448_036.aa.3	N	0 - 01-178;	18-28, 1.103; 39-66, 1.191; 69-78, 1.125; 139-159, 1.292; 85-100, 1.122;	MYRISTYL 168-173; MYRISTYL 41-46; PKC_PHOSPHO_SITE 95-97; MYRISTYL 153-158; MYRISTYL 131-136; PKC_PHOSPHO_SITE 48-50; MYRISTYL 123-128; MYRISTYL 47-52;	
DEX0448_036.orf.3	N	0 - 01-162;		MYRISTYL 107-112; PKC_PHOSPHO_SITE 32-34; MYRISTYL 115-120; PKC_PHOSPHO_SITE 79-81; MYRISTYL 137-142; MYRISTYL 152-157; MYRISTYL 27-32; MYRISTYL 31-36;	

169

			111-132, 1.258; 189-215, 1.126; 235-269, 1.226; 56-74, 1.194; 156-174, 1.168; 223-233, 1.072; 92-120, 1.188; 179-187, 1.124; 4-12, 1.15; 78- 86, 1.127; 122- 149, 1.229;		PKC_PHOSPHO_SITE 235-237; MYRISTYL 62-67; CK2_PHOSPHO_SITE 185-188; MYRISTYL 38-43; MYRISTYL 12-17; MYRISTYL 128-133; ASN_GLYCOSYLATION 268-271; MYRISTYL 28-33; MYRISTYL 26-31; MYRISTYL 15-20; MYRISTYL 115-120; MYRISTYL 34-39; MYRISTYL 265-270;	GLY_RICH 4-66; TMFOUR 94-120; TM4_1 105-127; TMFOUR 121-149; transmembrane4 45- 252; TM4_2 56-269; TMFOUR 240-266;
DEX0448_037.aa.1	N	3 - 01- 97;tm98- 120;i121- 126;tm127- 149;o150- 239;tm240- 262;i263-272;	91-105, 1.194; 123-151, 1.188; 28-36, 1.085; 266-300, 1.226; 74-84, 1.108; 44-50, 1.041; 55-66, 1.285; 109-117, 1.127; 210-218, 1.124; 220-246, 1.126; 254-264, 1.072; 153-180, 1.229; 187-205, 1.168; 4-22, 1.198;		MYRISTYL 93-98; MYRISTYL 146-151; MYRISTYL 13-18; MYRISTYL 3-8; PKC_PHOSPHO_SITE 266- 268; CK2_PHOSPHO_SITE 86-89; MYRISTYL 159- 164; ASN_GLYCOSYLATION 299-302; MYRISTYL 296-301; CK2_PHOSPHO_SITE 216-219;	TMFOUR 125-151; transmembrane4 90- 283; TMFOUR 271- 297; TMFOUR 152- 180; TM4_2 90-300; TM4_1 136-158;
DEX0448_037.aa.2	N	3 - 01- 128;tm129- 151;i152- 157;tm158- 180;o181- 270;tm271- 293;i294-303;	57-63, 1.102; 146-160, 1.137; 190-208, 1.221; 28-51, 1.074; 337-363, 1.126; 224-268, 1.188; 304-322, 1.168; 121-127, 1.055; 162-181, 1.262; 13-25, 1.066; 383-417, 1.226;		MYRISTYL 28-33; MYRISTYL 276-281; PKC_PHOSPHO_SITE 117-119; MYRISTYL 263- 268; PKC_PHOSPHO_SITE 383-385; MYRISTYL 229-234; MYRISTYL 226-231; MYRISTYL 137- 142; MYRISTYL 413-418; ASN_GLYCOSYLATION 416-419; CK2_PHOSPHO_SITE 150-153; CK2_PHOSPHO_SITE 333-336; MYRISTYL 9-14; CK2_PHOSPHO_SITE 161-164; MYRISTYL 6-11; ASN_GLYCOSYLATION 55-58;	TM4_1 253-275; TMFOUR 388-414; transmembrane4 225- 400; TM4_2 247-417; TMFOUR 269-297; TMFOUR 242-268;
DEX0448_037.aa.3	N	3 - 01- 245;tm246- 268;i269- 274;tm275- 297;o298- 387;tm388- 410;i411-420;				

			371-381, 1.072; 72-78, 1.077; 327-335, 1.124; 132-139, 1.09; 270-297, 1.229; 82-118, 1.227;			
DEX0448_037.orf.3	N	3 - o1- 72;tm73- 95;i96- 101;tm102- 124;o125- 214;tm215- 237;i238-247;	198-208, 1.072; 17-35, 1.221; 131-149, 1.168; 154-162, 1.124; 97-124, 1.229; 51-95, 1.188; 210-244, 1.226; 164-190, 1.126;	PKC_PHOSPHO_SITE 210-212; MYRISTYL 90-95; CK2_PHOSPHO_SITE 160-163; MYRISTYL 53-58; MYRISTYL 240-245; MYRISTYL 56-61; MYRISTYL 103-108; ASN_GLYCOSYLATION 243-246;	TM4_1 80-102; transmembrane4 52- 227; TM4_2 74-244; TMFOUR 215-241; TMFOUR 96-124; TMFOUR 69-95;	
DEX0448_037.aa.4	Y	4 - i1- 12;tm13- 35;o36- 61;tm62- 84;i85- 90;tm91- 113;o114- 203;tm204- 226;i227-236;	120-138, 1.168; 199-233, 1.226; 4-38, 1.194; 153-179, 1.126; 143-151, 1.124; 42-50, 1.127; 56-84, 1.188; 86-113, 1.229; 187-197, 1.072;	CK2_PHOSPHO_SITE 149-152; PKC_PHOSPHO_SITE 199-201; MYRISTYL 92-97; MYRISTYL 26-31; ASN_GLYCOSYLATION 232-235; MYRISTYL 229- 234; MYRISTYL 5-10; MYRISTYL 2-7; MYRISTYL 79-84;	TMFOUR 85-113; TM4_2 4-233; TM4_1 69-91; transmembrane4 9- 216; TMFOUR 204- 230; TMFOUR 12-35; TMFOUR 58-84;	
DEX0448_037.aa.5	Y	4 - i1- 12;tm13- 35;o36- 61;tm62- 84;i85- 90;tm91- 113;o114- 203;tm204- 226;i227-256;	42-50, 1.127; 227-239, 1.165; 199-225, 1.226; 120-138, 1.168; 86-113, 1.229; 143-151, 1.124; 4-38, 1.194; 187-197, 1.072; 56-84, 1.188; 153-179, 1.126;	PKC_PHOSPHO_SITE 239-241; MYRISTYL 92-97; CK2_PHOSPHO_SITE 149-152; MYRISTYL 79-84; PKC_PHOSPHO_SITE 199-201; MYRISTYL 26-31; MYRISTYL 2-7; MYRISTYL 5-10;	TMFOUR 204-230; TM4_1 69-91; TMFOUR 58-84; TMFOUR 12- 35; TMFOUR 85-113; transmembrane4 9- 216; TM4_2 4-226;	
DEX0448_037.aa.6	Y	3 - i1- 12;tm13- 35;o36-	4-38, 1.194; 187-197, 1.072; 56-84, 1.188;	PKC_PHOSPHO_SITE 199-201; MYRISTYL 26-31; MYRISTYL 79-84; MYRISTYL 92-97; PKC PHOSPHO SITE 208-210; MYRISTYL 5-10;	transmembrane4 9- 209; TMFOUR 85-113; TM4_2 4-210; TMFOUR	

			59;tm60-82;i83-88;tm89-111;o112-210;	120-138,1.168; 153-179,1.126; 86-113,1.229; 42-50,1.127; 143-151,1.124; 199-207,1.09;	MYRISTYL 2-7; CK2_PHOSPHO_SITE 149-152;	12-35; TMFOUR 58-84; TM4_1 69-91;
DEX0448_038.aa.1	N	0 - o1-58;		4-16,1.116; 21-30,1.102;	PKC_PHOSPHO_SITE 51-53;	SP_P04818_TYSY_HUMAN 1-58;
DEX0448_038.orf.1	N	0 - o1-91;		60-80,1.193; 42-49,1.089; 51-58,1.113;	PKC_PHOSPHO_SITE 34-36; PKC_PHOSPHO_SITE 40-42; PKC_PHOSPHO_SITE 29-31;	
DEX0448_039.aa.1	N	0 - o1-86;		67-79,1.076; 7-19,1.134; 36-49,1.1;	PKC_PHOSPHO_SITE 73-75; CK2_PHOSPHO_SITE 15-18; CK2_PHOSPHO_SITE 74-77; ASN_GLYCOSYLATION 57-60; PKC_PHOSPHO_SITE 15-17; CK2_PHOSPHO_SITE 47-50; CK2_PHOSPHO_SITE 4-7; CAMP_PHOSPHO_SITE 67-70;	
DEX0448_040.aa.1	N	0 - o1-107;		40-46,1.08; 59-78,1.174; 4-28,1.134; 52-57,1.069;	ASN_GLYCOSYLATION 96-99; PKC_PHOSPHO_SITE 25-27; PKC_PHOSPHO_SITE 36-38; MYRISTYL 49-54; CK2_PHOSPHO_SITE 36-39;	
DEX0448_040.orf.1	Y	0 - i1-90;		42-61,1.174; 23-29,1.08; 35-40,1.069; 4-15,1.154;	ASN_GLYCOSYLATION 79-82; CK2_PHOSPHO_SITE 19-22; MYRISTYL 32-37; PKC_PHOSPHO_SITE 19-21; MYRISTYL 5-10;	
DEX0448_040.aa.2	Y	0 - o1-193;		126-132,1.08; 59-79,1.208; 95-102,1.07; 104-114,1.156; 138-143,1.069; 20-33,1.175; 145-164,1.174; 83-88,1.089;	MYRISTYL 61-66; CK2_PHOSPHO_SITE 51-54; MYRISTYL 87-92; CK2_PHOSPHO_SITE 50-53; MYRISTYL 135-140; PKC_PHOSPHO_SITE 122-124; CK2_PHOSPHO_SITE 122-125; ASN_GLYCOSYLATION 182-185; MYRISTYL 7-12;	
DEX0448_040.orf.2	Y	0 - o1-189;		79-84,1.089; 16-29,1.175;	CK2_PHOSPHO_SITE 118-121; MYRISTYL 83-88; CK2_PHOSPHO_SITE 46-49; CK2 PHOSPHO SITE	

			134-139,1.069; 55-75,1.208; 122-128,1.08; 91-98,1.07; 100-110,1.156; 141-160,1.174;	47-50; PKC_PHOSPHO_SITE 118-120; MYRISTYL 57-62; MYRISTYL 131-136; ASN_GLYCOSYLATION 178-181;	
DEX0448_040.aa.3	Y	0 - 01-202;	104-111,1.07; 135-141,1.08; 113-123,1.156; 92-97,1.089; 154-173,1.174; 68-88,1.208; 29-42,1.175; 147-152,1.069;	CK2_PHOSPHO_SITE 131-134; CK2_PHOSPHO_SITE 59-62; ASN_GLYCOSYLATION 191-194; MYRISTYL 96-101; MYRISTYL 70-75; PKC_PHOSPHO_SITE 131-133; MYRISTYL 144-149; PKC_PHOSPHO_SITE 13-15; PKC_PHOSPHO_SITE 14-16; CK2_PHOSPHO_SITE 60-63;	
DEX0448_040.orf.3	Y	0 - 01-208;	141-147,1.08; 98-103,1.089; 74-94,1.208; 35-48,1.175; 153-158,1.069; 110-117,1.07; 119-129,1.156; 4-9,1.101; 160- 179,1.174;	MYRISTYL 8-13; CK2_PHOSPHO_SITE 24-27; PKC_PHOSPHO_SITE 137-139; MYRISTYL 12-17; CK2_PHOSPHO_SITE 65-68; MYRISTYL 76-81; MYRISTYL 150-155; MYRISTYL 16-21; ASN_GLYCOSYLATION 197-200; CK2_PHOSPHO_SITE 66-69; MYRISTYL 102-107; CK2_PHOSPHO_SITE 137-140;	
DEX0448_040.aa.4	Y	1 - i1- 11;tm12- 34;o35-189;	79-84,1.089; 16-29,1.182; 134-139,1.069; 141-160,1.174; 91-98,1.07; 55- 75,1.208; 122- 128,1.08; 100- 110,1.156;	PKC_PHOSPHO_SITE 118-120; MYRISTYL 83-88; ASN_GLYCOSYLATION 178-181; CK2_PHOSPHO_SITE 7-10; MYRISTYL 57-62; CK2_PHOSPHO_SITE 47-50; CK2_PHOSPHO_SITE 118-121; CK2_PHOSPHO_SITE 46-49; MYRISTYL 131-136;	
DEX0448_040.aa.5	N	0 - 01-171;	116-121,1.069; 104-110,1.08; 61-66,1.089; 123-142,1.174; 82-92,1.156;	CK2_PHOSPHO_SITE 29-32; MYRISTYL 39-44; MYRISTYL 12-17; CK2_PHOSPHO_SITE 100-103; ASN_GLYCOSYLATION 160-163; PKC_PHOSPHO_SITE 100-102; CK2_PHOSPHO_SITE 28-31; MYRISTYL 65-70; MYRISTYL 113-118;	

			37-57, 1.208; 12-17, 1.061; 73-80, 1.07;			
DEX0448_040.orf.5	N	0 - 01-171;			CK2 PHOSPHO SITE 29-32; MYRISTYL 65-70; ASN GLYCOSYLATION 160-163; PKC PHOSPHO_SITE 100-102; MYRISTYL 39-44; CK2 PHOSPHO_SITE 100-103; MYRISTYL 12-17; CK2 PHOSPHO_SITE 28-31; MYRISTYL 113-118;	
DEX0448_040.aa.6	N	0 - 01-156;	29-42, 1.16; 89- 95, 1.08; 108- 127, 1.174; 67- 77, 1.156; 58- 65, 1.07; 46- 51, 1.089; 101- 106, 1.069; 17- 24, 1.085;		PKC PHOSPHO SITE 28-30; MYRISTYL 50-55; PKC PHOSPHO_SITE 85-87; CK2 PHOSPHO_SITE 85-88; ASN GLYCOSYLATION 145-148; MYRISTYL 98-103;	
DEX0448_040.orf.6	N	0 - 01-156;	108-127, 1.174; 17-24, 1.08; 4- 9, 1.156; 67- 77, 1.156; 58- 65, 1.07; 101- 106, 1.069; 89- 95, 1.08; 46- 51, 1.089; 29- 42, 1.198;		MYRISTYL 98-103; PKC PHOSPHO_SITE 26-28; MYRISTYL 50-55; MYRISTYL 13-18; PKC PHOSPHO_SITE 85-87; ASN GLYCOSYLATION 145-148; CK2 PHOSPHO_SITE 85-88; AMIDATION 8-11;	
DEX0448_040.aa.7	N	0 - 01-176;	28-47, 1.174; 9- 15, 1.08; 21- 26, 1.069;		MYRISTYL 18-23; PKC PHOSPHO_SITE 5-7; CK2 PHOSPHO_SITE 5-8; ASN GLYCOSYLATION 65-68;	
DEX0448_040.orf.7	N	0 - 01-139;	68-85, 1.137; 127-136, 1.186; 31-42, 1.12; 97- 122, 1.182; 48- 62, 1.098; 12- 29, 1.167;		PKC PHOSPHO_SITE 44-46; PKC PHOSPHO_SITE 126-128; CK2 PHOSPHO_SITE 134-137; MYRISTYL 62-67; PKC PHOSPHO_SITE 25-27; MYRISTYL 4-9; CK2 PHOSPHO_SITE 11-14;	
DEX0448_041.aa.1	N	0 - 01-568;	349-355, 1.061;		CK2 PHOSPHO_SITE 356-359; MYRISTYL 300-	RRM RNP 1 429-436;

			<p>17-23, 1.06; 457-464, 1.094; 524-531, 1.068; 115-122, 1.095; 90-107, 1.114; 320-344, 1.172; 306-314, 1.081; 502-517, 1.085; 221-232, 1.182; 126-132, 1.03; 409-416, 1.032; 360-375, 1.098; 418-425, 1.091; 486-497, 1.133; 430-436, 1.057; 44-51, 1.122; 57-86, 1.155; 171-183, 1.064; 552-558, 1.032; 393-407, 1.102; 287-293, 1.019;</p>	<p>305; AMIDATION 122-125; MYRISTYL 475-480; ASN GLYCOSYLATION 317-320; CK2_PHOSPHO_SITE 319-322; CK2_PHOSPHO_SITE 206-209; CK2_PHOSPHO_SITE 386-389; PKC_PHOSPHO_SITE 69-71; MYRISTYL 133-138; CK2_PHOSPHO_SITE 145-148; CK2_PHOSPHO_SITE 438-441; CK2_PHOSPHO_SITE 42-45; CK2_PHOSPHO_SITE 496-499; AMIDATION 217- 220; CK2_PHOSPHO_SITE 452-455; AMIDATION 277-280; CK2_PHOSPHO_SITE 34-37; CK2_PHOSPHO_SITE 41-44; CK2_PHOSPHO_SITE 445-448; CK2_PHOSPHO_SITE 501-504; PKC_PHOSPHO_SITE 346-348; AMIDATION 107- 110; CK2_PHOSPHO_SITE 28-31; CK2_PHOSPHO_SITE 153-156; ASN GLYCOSYLATION 478-481; CK2_PHOSPHO_SITE 405-408; AMIDATION 52-55; PKC_PHOSPHO_SITE 386-388; MYRISTYL 111- 116; CK2_PHOSPHO_SITE 184-187; PKC_PHOSPHO_SITE 60-62; ASN GLYCOSYLATION 492-495; CK2_PHOSPHO_SITE 532-535; AMIDATION 85-88;</p>	<p>LYS_RICH 51-142; HISTONEH5 67-86; HISTONEH5 45-62; xrm 395-461; RRM_2 393-466; RRM_3 486- 560; RRM 487-556; xrm 309-378; GLU_RICH 143-275; RRM_RNP_1 523-530; RRM 308-379; RRM_1 307-383; ASP_RICH 144-261; RRM 394- 462; xrm 488-555; HISTONEH5 9-23;</p>
DEX0448_041.aa.2	N	0 - 01-520;	<p>409-416, 1.032; 17-23, 1.06; 287-293, 1.019; 320-344, 1.172; 349-355, 1.061; 502-517, 1.085; 44-51, 1.122; 126-132, 1.03; 486-497, 1.133; 171-183, 1.064; 418-425, 1.091; 115-122, 1.095; 393-407, 1.102; 360-375, 1.098; 221-232, 1.182;</p>	<p>CK2_PHOSPHO_SITE 496-499; MYRISTYL 111- 116; CK2_PHOSPHO_SITE 206-209; CK2_PHOSPHO_SITE 34-37; MYRISTYL 133-138; CK2_PHOSPHO_SITE 41-44; CK2_PHOSPHO_SITE 438-441; ASN GLYCOSYLATION 478-481; MYRISTYL 475-480; CK2_PHOSPHO_SITE 145- 148; PKC_PHOSPHO_SITE 69-71; CK2_PHOSPHO_SITE 445-448; PKC_PHOSPHO_SITE 60-62; ASN GLYCOSYLATION 492-495; CK2_PHOSPHO_SITE 501-504; CK2_PHOSPHO_SITE 28-31; MYRISTYL 300-305; CK2_PHOSPHO_SITE 356-359; AMIDATION 217-220; PKC_PHOSPHO_SITE 386-388; PKC_PHOSPHO_SITE 346-348; CK2_PHOSPHO_SITE 452-455; AMIDATION 107-110; ASN GLYCOSYLATION 317-</p>	<p>ASP_RICH 144-261; RRM 308-379; HISTONEH5 9-23; xrm 309-378; HISTONEH5 67-86; LYS_RICH 51- 142; RRM_2 393-466; GLU_RICH 143-275; xrm 395-461; RRM_1 307-383; RRM_RNP_1 429-436; HISTONEH5 45-62; RRM 394-462;</p>

175

		306-314, 1.081; 457-464, 1.094; 430-436, 1.057; 57-86, 1.155; 90-107, 1.114;	320; AMIDATION 277-280; CK2 PHOSPHO_SITE 386-389; CK2 PHOSPHO_SITE 405-408; AMIDATION 85-88; AMIDATION 52-55; CK2 PHOSPHO_SITE 153-156; CK2 PHOSPHO_SITE 184-187; CK2 PHOSPHO_SITE 319-322; AMIDATION 122-125; CK2 PHOSPHO_SITE 42-45;				
DEX0448_042.aa.1	N	433-454, 1.202; 116-122, 1.114; 27-33, 1.061; 168-174, 1.106; 486-494, 1.129; 36-55, 1.153; 10-25, 1.147; 498-504, 1.044; 143-156, 1.166; 158-165, 1.039; 234-249, 1.141; 331-352, 1.137; 468-478, 1.035; 372-383, 1.096; 95-101, 1.127; 268-277, 1.142; 222-230, 1.113; 60-65, 1.068; 251-264, 1.134; 283-303, 1.201; 75-87, 1.119; 395-407, 1.181;	MYRISTYL 77-82; CK2 PHOSPHO_SITE 497-500; CAMP PHOSPHO_SITE 249-252; MYRISTYL 143-148; AMIDATION 193-196; PKC PHOSPHO_SITE 200-202; PKC PHOSPHO_SITE 70-72; CK2 PHOSPHO_SITE 351-354; CK2 PHOSPHO_SITE 105-108; MYRISTYL 329-334; CK2 PHOSPHO_SITE 381-384; PKC PHOSPHO_SITE 471-473; MYRISTYL 484-489; CK2 PHOSPHO_SITE 206-209; MYRISTYL 439-444; ASN GLYCOSYLATION 230-233; CK2 PHOSPHO_SITE 163-166; MYRISTYL 183-188; CK2 PHOSPHO_SITE 164-167; ASN GLYCOSYLATION 103-106; CK2 PHOSPHO_SITE 200-203; ASN GLYCOSYLATION 426-429; CK2 PHOSPHO_SITE 410-413; PKC PHOSPHO_SITE 231-233; CK2 PHOSPHO_SITE 383-386; CAMP PHOSPHO_SITE 156-159; PKC PHOSPHO_SITE 247-249; MYRISTYL 435-440; MYRISTYL 112-117; MYRISTYL 322-327;	0 - 01-526;			CHAPERONIN60 107-134; CHAPERONIN60 423-444; cpn60_TCP1 47-526; CHAPERONIN60 51-77; CHAPERONIN60 375-400; TCOMPLEXTCP1 403-425; TCOMPLEXTCP1 105-124; TCOMPLEXTCP1 71-89; TCOMPLEXTCP1 436-448; CHAPERONINS_CPN60 430-441; CHAPERONIN60 292-315; TCOMPLEXTCP1 49-65;
DEX0448_043.aa.1	N	60-69, 1.144; 92-118, 1.156; 4-22, 1.179; 27-38, 1.084; 78-84, 1.06;	PKC PHOSPHO_SITE 70-72; CK2 PHOSPHO_SITE 39-42; PKC PHOSPHO_SITE 34-36; CK2 PHOSPHO_SITE 88-91; ASN GLYCOSYLATION 71-74; CK2 PHOSPHO_SITE 11-14; MYRISTYL 33-38; PKC PHOSPHO_SITE 88-90;	0 - 01-121;			sp_P07355_ANX2_HUMA N 41-114; ANNEXINII 11-23; annexin 43-111; ANNEXINII 96-112; ANNEXINII 56-78; ANNEXIN 56-78; ANNEXIN 96-112; ANX 59-111;

DEX0448_043.orf.1	N	0 - i1-71;	9-18,1.144; 41-68,1.156; 27-33,1.06;	PKC_PHOSPHO_SITE 37-39; CK2_PHOSPHO_SITE 2-5; CK2_PHOSPHO_SITE 37-40; PKC_PHOSPHO_SITE 19-21; ASN_GLYCOSYLATION 20-23;	ANNEXIN 45-61; ANX 8-60; ANNEXINII 5-27; ANNEXINII 45-61; ANNEXIN 5-27; SP_P07355_ANX2_HUMA N 7-63; annexin 2-60; ANNEXIN 8-60;
DEX0448_043.aa.2	N	0 - o1-242;	210-216,1.075; 234-239,1.075; 29-42,1.179; 111-145,1.156; 97-103,1.06; 18-25,1.062; 47-58,1.084; 7-14,1.138; 193-203,1.116; 179-185,1.11; 79-88,1.144;	CK2_PHOSPHO_SITE 6-9; PKC_PHOSPHO_SITE 26-28; CK2_PHOSPHO_SITE 163-166; MYRISTYL 53-58; PKC_PHOSPHO_SITE 107-109; CK2_PHOSPHO_SITE 194-197; CK2_PHOSPHO_SITE 173-176; CK2_PHOSPHO_SITE 146-149; AMIDATION 186-189; PKC_PHOSPHO_SITE 231-233; CK2_PHOSPHO_SITE 31-34; CAMP_PHOSPHO_SITE 218-221; PKC_PHOSPHO_SITE 89-91; ASN_GLYCOSYLATION 90-93; CK2_PHOSPHO_SITE 59-62; CK2_PHOSPHO_SITE 107-110; PKC_PHOSPHO_SITE 54-56;	ANNEXIN 75-97; ANX 142-184; ANNEXINII 75-97; ANX 78-130; SP_Q99KH3_Q99KH3_MO USE 142-187; SP_P07355_ANX2_HUMA N 61-133; ANNEXINII 115-131; annexin 63-130; annexin 135-184; SP_P07355_ANX2_HUMA N 203-235; ANNEXIN 78-130; ANNEXIN 115-131; ANNEXINII 31-43;
DEX0448_043.orf.2	N	0 - o1-342;	331-339,1.188; 224-244,1.161; 90-96,1.06; 288-299,1.146; 72-81,1.144; 315-324,1.131; 104-138,1.156; 40-51,1.084; 22-35,1.179; 186-196,1.116; 11-18,1.062; 257-277,1.177; 172-178,1.11;	MYRISTYL 46-51; AMIDATION 179-182; PKC_PHOSPHO_SITE 221-223; CK2_PHOSPHO_SITE 187-190; PKC_PHOSPHO_SITE 82-84; PKC_PHOSPHO_SITE 100-102; CK2_PHOSPHO_SITE 326-329; CK2_PHOSPHO_SITE 100-103; PKC_PHOSPHO_SITE 280-282; CK2_PHOSPHO_SITE 299-302; CK2_PHOSPHO_SITE 297-300; PKC_PHOSPHO_SITE 19-21; PKC_PHOSPHO_SITE 246-248; CK2_PHOSPHO_SITE 166-169; ASN_GLYCOSYLATION 83-86; PKC_PHOSPHO_SITE 47-49; CK2_PHOSPHO_SITE 139-142; CK2_PHOSPHO_SITE 24-27; CK2_PHOSPHO_SITE 52-55; CAMP_PHOSPHO_SITE 208-211; CK2_PHOSPHO_SITE 239-242; CK2_PHOSPHO_SITE	ANNEXINI 325-338; annexin 56-123; ANNEXINI 108-124; ANNEXINI 281-301; ANNEXINII 257-264; SP_P07355_ANX2_HUMA N 54-126; ANNEXINII 281-301; ANNEXINII 24-36; ANNEXINII 325-338; ANNEXINIV 108-124; ANNEXINIV 68-90; ANNEXINIV 281-307; ANX 71-123; ANNEXINIV 325-

177

				156-159;	338; ANNEXIN 68-90; ANNEXIN 285-337; ANNEXIN 108-124; annexin 270-337; ANNEXIN 71-123; ANX 285-337; ANNEXINIV 325-338; ANNEXIN 325-338; ANNEXIN 210-262; ANX 210- 262; ANNEXINIV 68- 90; ANNEXINIV 108- 124; sp_Q99KH3_Q99KH3_MO USE 272-340; sp_P07355_ANX2_HUMA N 196-265; ANNEXINIV 281-307; annexin 128-177; ANNEXINIV 43-53; sp_Q99KH3_Q99KH3_MO USE 135-180; ANNEXINII 201-227; ANNEXIN 281-301; annexin 194-262; ANNEXINI 68-90; ANX 135-177;
DEX0448_044.aa.1	Y	0 - 01-72;	4-27,1.179; 47- 52,1.073; 29- 35,1.043; 57- 67,1.134;		
DEX0448_044.orf.1	Y	0 - 01-103;	89-95,1.083; 30-43,1.171; 55-64,1.123; 73-81,1.126; 4- 19,1.14;	PKC PHOSPHO SITE 42-44; MYRISTYL 70-75; CAMP_PHOSPHO_SITE 37-40; PKC_PHOSPHO_SITE 65-67; MYRISTYL 9-14;	sp_Q9BYK1_Q9BYK1_HU MAN 23-70; RIBOSOMAL_S21E 33- 41; Ribosomal_S21e 23-103;
DEX0448_044.orf.2	Y	0 - 01-112;	73-95,1.144;	MYRISTYL 70-75; PKC PHOSPHO SITE 42-44;	SD O9BYK1 O9BYK1 HU

				30-43, 1.171; 100-109, 1.205; 55-64, 1.123; 4- 19, 1.14;	CAMP_PHOSPHO_SITE 37-40; PKC_PHOSPHO_SITE 65-67; MYRISTYL 9-14; MYRISTYL 96-101;	MAN 23-70; Ribosomal_S21e 23- 101; RIBOSOMAL_S21E 33-41;
DEX0448_044.aa.3	Y	0 - 01-87;		29-35, 1.08; 70- 84, 1.126; 19- 25, 1.06; 54- 61, 1.062; 5- 13, 1.128;	AMIDATION 64-67; MYRISTYL 7-12; PKC_PHOSPHO_SITE 33-35; MYRISTYL 36-41; ASN_GLYCOSYLATION 37-40; MYRISTYL 3-8; MYRISTYL 15-20; PKC_PHOSPHO_SITE 41-43; MYRISTYL 69-74; MYRISTYL 12-17;	sp_Q9BYK2_Q9BYK2_HU MAN 39-69; Ribosomal_S21e 35- 83;
DEX0448_044.orf.3	Y	0 - 01-86;		53-60, 1.062; 28-41, 1.171; 5- 17, 1.128; 69- 83, 1.126;	CAMP_PHOSPHO_SITE 35-38; AMIDATION 63-66; PKC_PHOSPHO_SITE 40-42; MYRISTYL 68-73; MYRISTYL 3-8; MYRISTYL 7-12;	sp_Q9BYK2_Q9BYK2_HU MAN 21-68; Ribosomal_S21e 21- 82; RIBOSOMAL_S21E 31-39;
DEX0448_044.aa.4	N	0 - 11-77;		63-69, 1.083; 47-55, 1.126; 29-38, 1.123;	ASN_GLYCOSYLATION 12-15; MYRISTYL 11-16; PKC_PHOSPHO_SITE 8-10; MYRISTYL 7-12; PKC_PHOSPHO_SITE 39-41; MYRISTYL 44-49; PKC_PHOSPHO_SITE 16-18;	Ribosomal_S21e 10- 77; sp_Q9BYK1_Q9BYK1_HU MAN 14-44;
DEX0448_044.orf.4	N	0 - 01-86;		15-26, 1.119; 72-78, 1.083; 56-64, 1.126; 4- 10, 1.034; 38- 47, 1.123;	PKC_PHOSPHO_SITE 25-27; MYRISTYL 53-58; CAMP_PHOSPHO_SITE 20-23; CK2_PHOSPHO_SITE 7-10; PKC_PHOSPHO_SITE 48-50;	Ribosomal_S21e 17- 86; sp_Q9BYK1_Q9BYK1_HU MAN 17-53;

Example 1b: Sequence Alignment Support

Alignments between previously identified sequences and splice variant sequences are performed to confirm unique portions of splice variant nucleic acid and amino acid sequences. The alignments are done using the Needle program in the European Molecular Biology Open Software Suite (EMBOSS) version 2.2.0 available at www.emboss.org from EMBnet (<http://www.embn.net.org>). Default settings are used unless otherwise noted. The Needle program in EMBOSS implements the Needleman-Wunsch algorithm. Needleman, S. B., Wunsch, C. D., *J. Mol. Biol.* 48:443-453 (1970).

It is well known to those skilled in the art that implication of alignment algorithms by various programs may result in minor changes in the generated output. These changes include but are not limited to: alignment scores (percent identity, similarity, and gap), display of nonaligned flanking sequence regions, and number assignment to residues. These minor changes in the output of an alignment do not alter the physical characteristics of the sequences or the differences between the sequences, e.g. regions of homology, insertions, or deletions.

Example 1c: RT-PCR Analysis

To detect the presence and tissue distribution of a particular splice variant Reverse Transcription-Polymerase Chain Reaction (RT-PCR) is performed using cDNA generated from a panel of tissue RNAs. See, e.g., Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press (1989) and; Kawasaki ES *et al.*, *PNAS* 85(15):5698 (1988). Total RNA is extracted from a variety of tissues and first strand cDNA is prepared with reverse transcriptase (RT). Each panel includes 23 cDNAs from five cancer types (lung, ovary, breast, colon, and prostate) and normal samples of testis, placenta and fetal brain. Each cancer set is composed of three cancer cDNAs from different donors and one normal pooled sample. Using a standard enzyme kit from BD Bioscience Clontech (Mountain View, CA), the target transcript is detected with sequence-specific primers designed to only amplify the particular splice variant. The PCR reaction is run on the GeneAmp PCR system 9700 (Applied Biosystem, Foster City, CA) thermocycler under optimal conditions. One of ordinary skill can design appropriate primers and determine optimal conditions. The amplified product is resolved on an agarose gel to detect a band of equivalent size to the predicted RT-PCR product. A band

indicated the presence of the splice variant in a sample. The relation of the amplified product to the splice variant was subsequently confirmed by DNA sequencing.

After subcloning, all positively screened clones are sequence verified. The DNA sequence verification results show the splice variant contains the predicted sequence
5 differences in comparison with the reference sequence.

Results for RT-PCR analysis include the Sequence DEX ID, Lead Name, Cancer Tissue(s) the transcript was detected in, Normal Tissue(s) the transcript was detected in, the predicted length of the RT-PCR product, and the Confirmed Length of the RT-PCR product.

10 RT-PCR results confirm the presence SEQ ID NO: 1-95 in biologic samples and distinguish between related transcripts.

Example 1d: Secretion Assay

To determine if a protein encoded by a splice variant is secreted from cells a secretion assay is preformed. A pcDNA3.1 clone containing the gene transcript which
15 encodes the variant protein is transfected into 293T cells using the Superfect transfection reagent (Qiagen, Valencia CA). Transfected cells are incubated for 28 hours before the media is collected and immediately spun down to remove any detached cells. The adherent cells are solubilized with lysis buffer (1% NP40, 10mM sodium phosphate pH7.0, and 0.15M NaCl). The lysed cells are collected and spun down and the
20 supernatant extracted as cell lysate. Western immunoblot is carried out in the following manner: 15µl of the cell lysate and media are run on 4-12% NuPage Bis-Tris gel (Invitrogen, Carlsbad CA), and blotted onto a PVDF membrane (Invitrogen, Carlsbad CA). The blot is incubated with a polyclonal primary antibody which binds to the variant protein (Imgenex, San Diego CA) and polyclonal goat anti-rabbit-peroxidase secondary
25 antibody (Sigma-Aldrich, St. Louis MO). The blot is developed with the ECL Plus chemiluminescent detection reagent (Amersham BioSciences, Piscataway NJ).

Secretion assay results are indicative of SEQ ID NO: 96-237 being a diagnostic marker and/or therapeutic target for cancer.

Example 2a: Gene Expression Analysis

30 *Custom Microarray Experiment - Cancer*

Custom oligonucleotide microarrays were provided by Agilent Technologies, Inc. (Palo Alto, CA). The microarrays were fabricated by Agilent using their technology for

the *in-situ* synthesis of 60mer oligonucleotides (Hughes, et al. 2001, Nature Biotechnology 19:342-347). The 60mer microarray probes were designed by Agilent, from gene sequences provided by diaDexus, using Agilent proprietary algorithms. Whenever possible two different 60mers were designed for each gene of interest.

5 All microarray experiments were two-color experiments and were preformed using Agilent-recommended protocols and reagents. Briefly, each microarray was hybridized with cRNAs synthesized from RNA (total RNA for ovarian and prostate, polyA+ RNA for lung, breast and colon samples), isolated from cancer and normal tissues, labeled with fluorescent dyes Cyanine3 (Cy3) or Cyanine5 (Cy5) (NEN Life Science Products, Inc.,
10 Boston, MA) using a linear amplification method (Agilent). In each experiment the experimental sample was RNA isolated from cancer tissue from a single individual and the reference sample was a pool of RNA isolated from normal tissues of the same organ as the cancerous tissue (*i.e.* normal ovarian tissue in experiments with ovarian cancer samples). Hybridizations were carried out at 60°C, overnight using Agilent *in-situ*
15 hybridization buffer. Following washing, arrays were scanned with a GenePix 4000B Microarray Scanner (Axon Instruments, Inc., Union City, CA). The resulting images were analyzed with GenePix Pro 3.0 Microarray Acquisition and Analysis Software (Axon).

Data normalization and expression profiling were done with Expressionist software from GeneData Inc. (Daly City, CA/Basel, Switzerland). Gene expression
20 analysis was performed using only experiments that met certain quality criteria. The quality criteria that experiments must meet are a combination of evaluations performed by the Expressionist software and evaluations performed manually using raw and normalized data. To evaluate raw data quality, detection limits (the mean signal for a replicated negative control + 2 Standard Deviations (SD)) for each channel were calculated. The
25 detection limit is a measure of non-specific hybridization. Acceptable detection limits were defined for each dye (<80 for Cy5 and <150 for Cy3). Arrays with poor detection limits in one or both channels were not analyzed and the experiments were repeated. To evaluate normalized data quality, positive control elements included in the array were utilized. These array features should have a mean ratio of 1 (no differential expression).
30 If these features have a mean ratio of greater than 1.5-fold up or down, the experiments were not analyzed further and were repeated. In addition to traditional scatter plots demonstrating the distribution of signal in each experiment, the Expressionist software also has minimum thresholding criteria that employ user defined parameters to identify

quality data. These thresholds include two distinct quality measurements: 1) minimum area percentage, which is a measure of the integrity of each spot and 2) signal to noise ratio, which ensures that the signal being measured is significantly above any background (nonspecific) signal present. Only those features that met the threshold criteria were
5 included in the filtering and analyses carried out by Expressionist. The thresholding settings employed require a minimum area percentage of 60% [(% pixels > background + 2SD)-(% pixels saturated)], and a minimum signal to noise ratio of 2.0 in both channels. By these criteria, very low expressors, saturated features and spots with abnormally high local background were not included in analysis.

10 Relative expression data was collected from Expressionist based on filtering and clustering analyses. Up-regulated genes were identified using criteria for the percentage of experiments in which the gene is up-regulated by at least 2-fold. In general, up-regulation in ~30% of samples tested was used as a cutoff for filtering.

Two microarray experiments were performed for each normal and cancer tissue
15 pair. The tissue specific Array Chip for each cancer tissue is a unique microarray specific to that tissue and cancer. The Multi-Cancer Array Chip is a universal microarray that was hybridized with samples from each of the cancers (ovarian, breast, colon, lung, and prostate). See the description below for the experiments specific to the different cancers.

Microarray Experiments and Data Tables

20 COLON CANCER CHIPS

For colon cancer two different chip designs were evaluated with overlapping sets of a total of 38 samples, comparing the expression patterns of colon cancer derived polyA+ RNA to polyA+ RNA isolated from a pool of 7 normal colon tissues. For the Colon Array Chip all 38 samples (23 Ascending colon carcinomas and 15 Rectosigmoidal
25 carcinomas including: 5 stage I cancers, 15 stage II cancers, 15 stage III and 2 stage IV cancers, as well as 28 Grade 1/2 and 10 Grade 3 cancers) were analyzed. The histopathologic grades for cancer are classified as follows: GX, cannot be assessed; G1, well differentiated; G2, Moderately differentiated; G3, poorly differentiated; and G4, undifferentiated. AJCC Cancer Staging Handbook, 5th Edition, 1998, page 9. For the
30 Colon Array Chip analysis, samples were further divided into groups based on the expression pattern of the known colon cancer associated gene Thymidilate Synthase (TS) (13 TS up 25 TS not up). The association of TS with advanced colorectal cancer is well

documented. Paradiso *et al.*, *Br J Cancer* 82(3):560-7 (2000); Etienne *et al.*, *J Clin Oncol.* 20(12):2832-43 (2002); Aschele *et al.* *Clin Cancer Res.* 6(12):4797-802 (2000). For the Multi-Cancer Array Chip a subset of 27 of these samples (14 Ascending colon carcinomas and 13 Rectosigmoidal carcinomas including: 3 stage I cancers, 9 stage II cancers, 13 stage III and 2 stage IV cancers) were assessed.

The results for the statistically significant up-regulated genes on the Colon Array Chip are shown in Table 1 and 2. The results for the statistically significant up-regulated genes on the Multi-Cancer Array Chip are shown in Table 3.

The first two columns of each table contain information about the sequence itself (Seq ID, Oligo Name), the next columns show the results obtained for all ("ALL") the colon samples, ascending colon carcinomas ("ASC"), Rectosigmoidal carcinomas ("RS"), cancers corresponding to stages I and II ("ST1,2"), stages III and IV ("ST3,4"), grades 1 and 2 ("GR1,2"), grade 3 ("GR3"), cancers exhibiting up-regulation of the TS gene ("TSup") or those not exhibiting up-regulation of the TS gene ("NOT TSup"). '%up' indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed n=38 for the Colon Array Chip (n=27 for the Multi-Cancer Array Chip), '%valid up' indicates the percentage of experiments with valid expression values in which up-regulation of at least 2-fold was observed.

Table 1.

DEX ID	Oligo Name	Cln ALL %up n=38	Cln ALL %valid up n=38	Cln ASC %up n=23	Cln ASC %valid up n=23	Cln RS %up n=15	Cln RS %valid up n=15	Cln ST1,2 %up n=20	Cln ST1,2 %valid up n=20	Cln ST3,4 %up n=18	Cln ST3,4 %valid up n=18
DEX0448 001.nt.1	34940.0	23.7	23.7	30.4	30.4	13.3	13.3	25.0	25.0	22.2	22.2
DEX0448 002.nt.1	39957.0	52.6	52.6	56.5	56.5	46.7	46.7	70.0	70.0	33.3	33.3
DEX0448 002.nt.1	39958.0	44.7	44.7	43.5	43.5	46.7	46.7	60.0	60.0	27.8	27.8
DEX0448 003.nt.1	32057.0	18.4	18.4	26.1	26.1	6.7	6.7	15.0	15.0	22.2	22.2
DEX0448 004.nt.1	41210.0	21.1	21.1	30.4	30.4	6.7	6.7	15.0	15.0	27.8	27.8
DEX0448 005.nt.2	36243.0	34.2	34.2	43.5	43.5	20.0	20.0	40.0	40.0	27.8	27.8
DEX0448 006.nt.1	8410.0	13.2	13.2	13.0	13.0	13.3	13.3	15.0	15.0	11.1	11.1
DEX0448 007.nt.1	26449.0	10.5	10.8	13.0	13.0	6.7	7.1	5.0	5.0	16.7	17.6
DEX0448 008.nt.1	32851.0	28.9	28.9	30.4	30.4	26.7	26.7	35.0	35.0	22.2	22.2
DEX0448 009.nt.1	34680.0	55.3	61.8	56.5	59.1	53.3	66.7	45.0	56.2	66.7	66.7
DEX0448 009.nt.1	40809.0	57.9	57.9	56.5	56.5	60.0	60.0	55.0	55.0	61.1	61.1
DEX0448 009.nt.1	40810.0	57.9	57.9	56.5	56.5	60.0	60.0	55.0	55.0	61.1	61.1
DEX0448 010.nt.1	28423.0	18.4	18.9	30.4	30.4	0.0	0.0	20.0	20.0	16.7	17.6
DEX0448 011.nt.1	8312.0	2.6	2.8	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.9
DEX0448 013.nt.1	16006.0	18.4	18.4	26.1	26.1	6.7	6.7	5.0	5.0	33.3	33.3
DEX0448 014.nt.1	37376.0	23.7	23.7	13.0	13.0	40.0	40.0	20.0	20.0	27.8	27.8
DEX0448 014.nt.1	37378.0	21.1	21.1	13.0	13.0	33.3	33.3	20.0	20.0	22.2	22.2
DEX0448 015.nt.1	33348.0	34.2	34.2	34.8	34.8	33.3	33.3	30.0	30.0	38.9	38.9
DEX0448 015.nt.1	38996.0	28.9	29.7	26.1	27.3	33.3	33.3	25.0	25.0	33.3	35.3

DEX0448	017.nt.1	36179.0	15.8	16.7	8.7	9.5	26.7	26.7	15.0	16.7	16.7	16.7
DEX0448	017.nt.1	36180.0	10.5	11.1	8.7	9.5	13.3	13.3	15.0	16.7	5.6	5.6
DEX0448	017.nt.1	36181.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	017.nt.1	36182.0	2.6	3.2	4.3	5.6	0.0	0.0	0.0	0.0	5.6	6.2
DEX0448	017.nt.1	36184.0	36.8	36.8	43.5	43.5	26.7	26.7	45.0	45.0	27.8	27.8
DEX0448	017.nt.1	37227.0	5.3	6.5	4.3	5.9	6.7	7.1	5.0	6.2	5.6	6.7
DEX0448	017.nt.1	37228.0	2.6	3.1	4.3	5.6	0.0	0.0	0.0	0.0	5.6	6.2
DEX0448	017.nt.1	37239.0	28.9	28.9	34.8	34.8	20.0	20.0	30.0	30.0	27.8	27.8
DEX0448	017.nt.1	37240.0	34.2	35.1	39.1	39.1	26.7	28.6	35.0	35.0	33.3	35.3
DEX0448	017.nt.1	37895.0	5.3	10.5	0.0	0.0	13.3	33.3	0.0	0.0	11.1	20.0
DEX0448	018.nt.1	41823.0	23.7	23.7	26.1	26.1	20.0	20.0	25.0	25.0	22.2	22.2
DEX0448	018.nt.1	41824.0	15.8	15.8	21.7	21.7	6.7	6.7	15.0	15.0	16.7	16.7
DEX0448	020.nt.1	39655.0	44.7	44.7	56.5	56.5	26.7	26.7	40.0	40.0	50.0	50.0
DEX0448	020.nt.1	39656.0	26.3	26.3	30.4	30.4	20.0	20.0	25.0	25.0	27.8	27.8
DEX0448	021.nt.1	30819.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.1	30870.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.1	30931.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.1	31146.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.1	31154.0	7.9	7.9	8.7	8.7	6.7	6.7	0.0	0.0	16.7	16.7
DEX0448	021.nt.1	31155.0	7.9	8.6	8.7	10.0	6.7	6.7	5.0	5.6	11.1	11.8
DEX0448	021.nt.1	31157.0	2.6	2.6	0.0	0.0	6.7	6.7	0.0	0.0	5.6	5.6
DEX0448	021.nt.1	35218.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.2	30819.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.2	30870.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.2	30931.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.2	31146.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.2	31154.0	7.9	7.9	8.7	8.7	6.7	6.7	0.0	0.0	16.7	16.7
DEX0448	021.nt.2	31155.0	7.9	8.6	8.7	10.0	6.7	6.7	5.0	5.6	11.1	11.8
DEX0448	021.nt.2	31157.0	2.6	2.6	0.0	0.0	6.7	6.7	0.0	0.0	5.6	5.6
DEX0448	021.nt.2	35218.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.3	30819.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.3	30870.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.3	31154.0	7.9	7.9	8.7	8.7	6.7	6.7	0.0	0.0	16.7	16.7
DEX0448	021.nt.3	31157.0	2.6	2.6	0.0	0.0	6.7	6.7	0.0	0.0	5.6	5.6
DEX0448	021.nt.3	35218.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	022.nt.1	36878.0	36.8	37.8	47.8	47.8	20.0	21.4	25.0	26.3	50.0	50.0
DEX0448	022.nt.1	37415.0	52.6	52.6	69.6	69.6	26.7	26.7	50.0	50.0	55.6	55.6
DEX0448	022.nt.2	36878.0	36.8	37.8	47.8	47.8	20.0	21.4	25.0	26.3	50.0	50.0
DEX0448	022.nt.2	37415.0	52.6	52.6	69.6	69.6	26.7	26.7	50.0	50.0	55.6	55.6
DEX0448	023.nt.1	22297.0	21.1	21.1	30.4	30.4	6.7	6.7	20.0	20.0	22.2	22.2
DEX0448	024.nt.1	19607.0	36.8	36.8	43.5	43.5	26.7	26.7	30.0	30.0	44.4	44.4
DEX0448	025.nt.1	40033.0	81.6	81.6	82.6	82.6	80.0	80.0	80.0	80.0	83.3	83.3
DEX0448	025.nt.1	40034.0	78.9	78.9	82.6	82.6	73.3	73.3	75.0	75.0	83.3	83.3
DEX0448	026.nt.1	41210.0	21.1	21.1	30.4	30.4	6.7	6.7	15.0	15.0	27.8	27.8
DEX0448	026.nt.1	41284.0	5.3	5.3	8.7	8.7	0.0	0.0	5.0	5.0	5.6	5.6
DEX0448	026.nt.2	41210.0	21.1	21.1	30.4	30.4	6.7	6.7	15.0	15.0	27.8	27.8
DEX0448	026.nt.2	41284.0	5.3	5.3	8.7	8.7	0.0	0.0	5.0	5.0	5.6	5.6
DEX0448	027.nt.1	9622.0	15.8	15.8	21.7	21.7	6.7	6.7	15.0	15.0	16.7	16.7
DEX0448	027.nt.1	9623.0	15.8	15.8	21.7	21.7	6.7	6.7	15.0	15.0	16.7	16.7
DEX0448	027.nt.2	9622.0	15.8	15.8	21.7	21.7	6.7	6.7	15.0	15.0	16.7	16.7
DEX0448	027.nt.2	9623.0	15.8	15.8	21.7	21.7	6.7	6.7	15.0	15.0	16.7	16.7
DEX0448	027.nt.3	9622.0	15.8	15.8	21.7	21.7	6.7	6.7	15.0	15.0	16.7	16.7
DEX0448	027.nt.3	9623.0	15.8	15.8	21.7	21.7	6.7	6.7	15.0	15.0	16.7	16.7
DEX0448	027.nt.4	9622.0	15.8	15.8	21.7	21.7	6.7	6.7	15.0	15.0	16.7	16.7
DEX0448	027.nt.4	9623.0	15.8	15.8	21.7	21.7	6.7	6.7	15.0	15.0	16.7	16.7
DEX0448	027.nt.5	9622.0	15.8	15.8	21.7	21.7	6.7	6.7	15.0	15.0	16.7	16.7
DEX0448	027.nt.5	9623.0	15.8	15.8	21.7	21.7	6.7	6.7	15.0	15.0	16.7	16.7

DEX0448 027.nt.6	9622.0	15.8	15.8	21.7	21.7	6.7	6.7	15.0	15.0	16.7	16.7
DEX0448 027.nt.6	9623.0	15.8	15.8	21.7	21.7	6.7	6.7	15.0	15.0	16.7	16.7
DEX0448 029.nt.1	28733.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.1	28734.0	13.2	13.2	13.0	13.0	13.3	13.3	5.0	5.0	22.2	22.2
DEX0448 029.nt.1	38381.0	2.6	2.9	0.0	0.0	6.7	7.7	0.0	0.0	5.6	6.7
DEX0448 029.nt.1	38382.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.1	38383.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 029.nt.1	38384.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 029.nt.2	28733.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.2	28734.0	13.2	13.2	13.0	13.0	13.3	13.3	5.0	5.0	22.2	22.2
DEX0448 029.nt.2	38381.0	2.6	2.9	0.0	0.0	6.7	7.7	0.0	0.0	5.6	6.7
DEX0448 029.nt.2	38382.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.2	38383.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 029.nt.2	38384.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 029.nt.3	28733.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.3	28734.0	13.2	13.2	13.0	13.0	13.3	13.3	5.0	5.0	22.2	22.2
DEX0448 029.nt.3	38381.0	2.6	2.9	0.0	0.0	6.7	7.7	0.0	0.0	5.6	6.7
DEX0448 029.nt.3	38382.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.3	38383.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 029.nt.3	38384.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 029.nt.4	28733.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.4	28734.0	13.2	13.2	13.0	13.0	13.3	13.3	5.0	5.0	22.2	22.2
DEX0448 029.nt.4	38381.0	2.6	2.9	0.0	0.0	6.7	7.7	0.0	0.0	5.6	6.7
DEX0448 029.nt.4	38382.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.4	38383.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 029.nt.4	38384.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 029.nt.5	28733.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.5	28734.0	13.2	13.2	13.0	13.0	13.3	13.3	5.0	5.0	22.2	22.2
DEX0448 029.nt.5	38381.0	2.6	2.9	0.0	0.0	6.7	7.7	0.0	0.0	5.6	6.7
DEX0448 029.nt.5	38382.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.5	38383.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 029.nt.5	38384.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 029.nt.6	28733.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.6	28734.0	13.2	13.2	13.0	13.0	13.3	13.3	5.0	5.0	22.2	22.2
DEX0448 029.nt.6	38381.0	2.6	2.9	0.0	0.0	6.7	7.7	0.0	0.0	5.6	6.7
DEX0448 029.nt.6	38382.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.6	38383.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 029.nt.6	38384.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 029.nt.7	28733.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.7	28734.0	13.2	13.2	13.0	13.0	13.3	13.3	5.0	5.0	22.2	22.2
DEX0448 029.nt.7	38381.0	2.6	2.9	0.0	0.0	6.7	7.7	0.0	0.0	5.6	6.7
DEX0448 029.nt.7	38382.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 029.nt.7	38383.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 029.nt.7	38384.0	2.6	2.6	4.3	4.3	0.0	0.0	0.0	0.0	5.6	5.6
DEX0448 030.nt.1	29603.0	5.3	5.3	8.7	8.7	0.0	0.0	5.0	5.0	5.6	5.6
DEX0448 030.nt.1	29604.0	7.9	8.1	13.0	13.0	0.0	0.0	10.0	10.0	5.6	5.9
DEX0448 030.nt.1	40867.0	10.5	11.4	17.4	19.0	0.0	0.0	5.0	5.3	16.7	18.8
DEX0448 030.nt.1	40868.0	13.2	13.2	21.7	21.7	0.0	0.0	5.0	5.0	22.2	22.2
DEX0448 032.nt.1	36179.0	15.8	16.7	8.7	9.5	26.7	26.7	15.0	16.7	16.7	16.7
DEX0448 032.nt.1	36180.0	10.5	11.1	8.7	9.5	13.3	13.3	15.0	16.7	5.6	5.6
DEX0448 032.nt.1	36182.0	2.6	3.2	4.3	5.6	0.0	0.0	0.0	0.0	5.6	6.2
DEX0448 032.nt.1	36184.0	36.8	36.8	43.5	43.5	26.7	26.7	45.0	45.0	27.8	27.8
DEX0448 032.nt.1	37227.0	5.3	6.5	4.3	5.9	6.7	7.1	5.0	6.2	5.6	6.7
DEX0448 032.nt.1	37228.0	2.6	3.1	4.3	5.6	0.0	0.0	0.0	0.0	5.6	6.2
DEX0448 032.nt.1	37239.0	28.9	28.9	34.8	34.8	20.0	20.0	30.0	30.0	27.8	27.8
DEX0448 032.nt.1	37240.0	34.2	35.1	39.1	39.1	26.7	28.6	35.0	35.0	33.3	35.3
DEX0448 035.nt.1	23862.0	18.4	18.4	13.0	13.0	26.7	26.7	10.0	10.0	27.8	27.8

DEX0448	035.nt.1	23863.0	26.3	26.3	26.1	26.1	26.7	26.7	20.0	20.0	33.3	33.3
DEX0448	035.nt.2	23862.0	18.4	18.4	13.0	13.0	26.7	26.7	10.0	10.0	27.8	27.8
DEX0448	035.nt.2	23863.0	26.3	26.3	26.1	26.1	26.7	26.7	20.0	20.0	33.3	33.3
DEX0448	035.nt.3	23862.0	18.4	18.4	13.0	13.0	26.7	26.7	10.0	10.0	27.8	27.8
DEX0448	035.nt.3	23863.0	26.3	26.3	26.1	26.1	26.7	26.7	20.0	20.0	33.3	33.3
DEX0448	035.nt.4	23862.0	18.4	18.4	13.0	13.0	26.7	26.7	10.0	10.0	27.8	27.8
DEX0448	035.nt.4	23863.0	26.3	26.3	26.1	26.1	26.7	26.7	20.0	20.0	33.3	33.3
DEX0448	036.nt.1	29532.0	26.3	26.3	34.8	34.8	13.3	13.3	20.0	20.0	33.3	33.3
DEX0448	036.nt.1	29533.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	036.nt.1	29534.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	036.nt.1	29540.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	036.nt.1	36811.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	036.nt.1	36812.0	2.6	3.2	0.0	0.0	6.7	8.3	0.0	0.0	5.6	6.7
DEX0448	036.nt.2	29532.0	26.3	26.3	34.8	34.8	13.3	13.3	20.0	20.0	33.3	33.3
DEX0448	036.nt.2	29539.0	2.6	2.7	0.0	0.0	6.7	6.7	0.0	0.0	5.6	5.9
DEX0448	036.nt.2	29540.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	036.nt.3	29532.0	26.3	26.3	34.8	34.8	13.3	13.3	20.0	20.0	33.3	33.3
DEX0448	036.nt.3	29539.0	2.6	2.7	0.0	0.0	6.7	6.7	0.0	0.0	5.6	5.9
DEX0448	036.nt.3	29540.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	037.nt.1	36341.0	10.5	10.5	13.0	13.0	6.7	6.7	5.0	5.0	16.7	16.7
DEX0448	037.nt.1	36342.0	7.9	8.3	8.7	8.7	6.7	7.7	5.0	5.3	11.1	11.8
DEX0448	037.nt.2	36341.0	10.5	10.5	13.0	13.0	6.7	6.7	5.0	5.0	16.7	16.7
DEX0448	037.nt.2	36342.0	7.9	8.3	8.7	8.7	6.7	7.7	5.0	5.3	11.1	11.8
DEX0448	037.nt.2	36341.0	10.5	10.5	13.0	13.0	6.7	6.7	5.0	5.0	16.7	16.7
DEX0448	037.nt.3	36342.0	7.9	8.3	8.7	8.7	6.7	7.7	5.0	5.3	11.1	11.8
DEX0448	037.nt.4	36341.0	10.5	10.5	13.0	13.0	6.7	6.7	5.0	5.0	16.7	16.7
DEX0448	037.nt.4	36342.0	7.9	8.3	8.7	8.7	6.7	7.7	5.0	5.3	11.1	11.8
DEX0448	037.nt.5	36341.0	10.5	10.5	13.0	13.0	6.7	6.7	5.0	5.0	16.7	16.7
DEX0448	037.nt.5	36342.0	7.9	8.3	8.7	8.7	6.7	7.7	5.0	5.3	11.1	11.8
DEX0448	037.nt.6	36341.0	10.5	10.5	13.0	13.0	6.7	6.7	5.0	5.0	16.7	16.7
DEX0448	037.nt.6	36342.0	7.9	8.3	8.7	8.7	6.7	7.7	5.0	5.3	11.1	11.8
DEX0448	038.nt.1	20895.0	28.9	29.7	34.8	36.4	20.0	20.0	30.0	31.6	27.8	27.8
DEX0448	038.nt.1	20896.0	31.6	31.6	39.1	39.1	20.0	20.0	35.0	35.0	27.8	27.8
DEX0448	039.nt.1	38855.0	34.2	34.2	34.8	34.8	33.3	33.3	35.0	35.0	33.3	33.3
DEX0448	040.nt.1	31347.0	31.6	32.4	34.8	34.8	26.7	28.6	40.0	40.0	22.2	23.5
DEX0448	040.nt.1	40755.0	28.9	28.9	34.8	34.8	20.0	20.0	35.0	35.0	22.2	22.2
DEX0448	040.nt.1	40756.0	31.6	31.6	34.8	34.8	26.7	26.7	40.0	40.0	22.2	22.2
DEX0448	040.nt.2	31346.0	36.8	36.8	39.1	39.1	33.3	33.3	40.0	40.0	33.3	33.3
DEX0448	040.nt.2	31347.0	31.6	32.4	34.8	34.8	26.7	28.6	40.0	40.0	22.2	23.5
DEX0448	040.nt.2	40755.0	28.9	28.9	34.8	34.8	20.0	20.0	35.0	35.0	22.2	22.2
DEX0448	040.nt.2	40756.0	31.6	31.6	34.8	34.8	26.7	26.7	40.0	40.0	22.2	22.2
DEX0448	040.nt.3	31346.0	36.8	36.8	39.1	39.1	33.3	33.3	40.0	40.0	33.3	33.3
DEX0448	040.nt.3	31347.0	31.6	32.4	34.8	34.8	26.7	28.6	40.0	40.0	22.2	23.5
DEX0448	040.nt.3	40755.0	28.9	28.9	34.8	34.8	20.0	20.0	35.0	35.0	22.2	22.2
DEX0448	040.nt.3	40756.0	31.6	31.6	34.8	34.8	26.7	26.7	40.0	40.0	22.2	22.2
DEX0448	040.nt.4	31346.0	36.8	36.8	39.1	39.1	33.3	33.3	40.0	40.0	33.3	33.3
DEX0448	040.nt.4	31347.0	31.6	32.4	34.8	34.8	26.7	28.6	40.0	40.0	22.2	23.5
DEX0448	040.nt.4	40755.0	28.9	28.9	34.8	34.8	20.0	20.0	35.0	35.0	22.2	22.2
DEX0448	040.nt.4	40756.0	31.6	31.6	34.8	34.8	26.7	26.7	40.0	40.0	22.2	22.2
DEX0448	040.nt.5	31346.0	36.8	36.8	39.1	39.1	33.3	33.3	40.0	40.0	33.3	33.3
DEX0448	040.nt.5	31347.0	31.6	32.4	34.8	34.8	26.7	28.6	40.0	40.0	22.2	23.5
DEX0448	040.nt.5	40755.0	28.9	28.9	34.8	34.8	20.0	20.0	35.0	35.0	22.2	22.2
DEX0448	040.nt.5	40756.0	31.6	31.6	34.8	34.8	26.7	26.7	40.0	40.0	22.2	22.2
DEX0448	040.nt.6	31346.0	36.8	36.8	39.1	39.1	33.3	33.3	40.0	40.0	33.3	33.3
DEX0448	040.nt.6	31347.0	31.6	32.4	34.8	34.8	26.7	28.6	40.0	40.0	22.2	23.5
DEX0448	040.nt.6	35943.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	040.nt.6	35944.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

DEX0448	040.nt.6	40755.0	28.9	28.9	34.8	34.8	20.0	20.0	35.0	35.0	22.2	22.2
DEX0448	040.nt.6	40756.0	31.6	31.6	34.8	34.8	26.7	26.7	40.0	40.0	22.2	22.2
DEX0448	040.nt.7	31346.0	36.8	36.8	39.1	39.1	33.3	33.3	40.0	40.0	33.3	33.3
DEX0448	040.nt.7	31347.0	31.6	32.4	34.8	34.8	26.7	28.6	40.0	40.0	22.2	23.5
DEX0448	040.nt.7	40755.0	28.9	28.9	34.8	34.8	20.0	20.0	35.0	35.0	22.2	22.2
DEX0448	040.nt.7	40756.0	31.6	31.6	34.8	34.8	26.7	26.7	40.0	40.0	22.2	22.2
DEX0448	041.nt.1	32216.0	52.6	52.6	52.2	52.2	53.3	53.3	55.0	55.0	50.0	50.0
DEX0448	042.nt.1	29231.0	15.8	16.2	13.0	13.0	20.0	21.4	20.0	21.1	11.1	11.1
DEX0448	042.nt.1	29232.0	15.8	15.8	13.0	13.0	20.0	20.0	20.0	20.0	11.1	11.1
DEX0448	042.nt.1	29271.0	23.7	23.7	21.7	21.7	26.7	26.7	30.0	30.0	16.7	16.7
DEX0448	042.nt.1	29272.0	26.3	26.3	21.7	21.7	33.3	33.3	30.0	30.0	22.2	22.2
DEX0448	042.nt.1	29290.0	31.6	31.6	30.4	30.4	33.3	33.3	45.0	45.0	16.7	16.7
DEX0448	043.nt.1	9043.0	18.4	18.4	17.7	21.7	13.3	13.3	15.0	15.0	22.2	22.2
DEX0448	044.nt.1	36803.0	28.9	28.9	17.4	17.4	46.7	46.7	25.0	25.0	33.3	33.3

Table 2.

[illegible]

DEX0448	021.nt.1	30870.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.1	30931.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.1	31146.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.1	31154.0	0.0	0.0	30.0	30.0	15.4	15.4	4.0
DEX0448	021.nt.1	31155.0	0.0	0.0	30.0	33.3	23.1	23.1	0.0
DEX0448	021.nt.1	31157.0	0.0	0.0	10.0	10.0	7.7	7.7	0.0
DEX0448	021.nt.1	35218.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.2	30819.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.2	30870.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.2	30931.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.2	31146.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.2	31154.0	0.0	0.0	30.0	30.0	15.4	15.4	4.0
DEX0448	021.nt.2	31155.0	0.0	0.0	30.0	33.3	23.1	23.1	0.0
DEX0448	021.nt.2	31157.0	0.0	0.0	10.0	10.0	7.7	7.7	0.0
DEX0448	021.nt.2	35218.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.3	30819.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.3	30870.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	021.nt.3	31154.0	0.0	0.0	30.0	30.0	15.4	15.4	4.0
DEX0448	021.nt.3	31157.0	0.0	0.0	10.0	10.0	7.7	7.7	0.0
DEX0448	021.nt.3	35218.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	022.nt.1	36878.0	39.3	40.7	30.0	30.0	38.5	38.5	36.0
DEX0448	022.nt.1	37415.0	46.4	46.4	70.0	70.0	46.2	46.2	56.0
DEX0448	022.nt.2	36878.0	39.3	40.7	30.0	30.0	38.5	38.5	36.0
DEX0448	022.nt.2	37415.0	46.4	46.4	70.0	70.0	46.2	46.2	56.0
DEX0448	023.nt.1	22297.0	17.9	17.9	30.0	30.0	38.5	38.5	12.0
DEX0448	024.nt.1	19607.0	32.1	32.1	50.0	50.0	46.2	46.2	32.0
DEX0448	025.nt.1	40033.0	85.7	85.7	70.0	70.0	76.9	76.9	84.0
DEX0448	025.nt.1	40034.0	82.1	82.1	70.0	70.0	76.9	76.9	80.0
DEX0448	026.nt.1	41210.0	10.7	10.7	50.0	50.0	38.5	38.5	12.0
DEX0448	026.nt.1	41284.0	7.1	7.1	0.0	0.0	0.0	0.0	8.0
DEX0448	026.nt.2	41210.0	10.7	10.7	50.0	50.0	38.5	38.5	12.0
DEX0448	026.nt.2	41284.0	7.1	7.1	0.0	0.0	0.0	0.0	8.0
DEX0448	027.nt.1	9622.0	14.3	14.3	20.0	20.0	30.8	30.8	8.0
DEX0448	027.nt.1	9623.0	14.3	14.3	20.0	20.0	30.8	30.8	8.0
DEX0448	027.nt.2	9622.0	14.3	14.3	20.0	20.0	30.8	30.8	8.0
DEX0448	027.nt.2	9623.0	14.3	14.3	20.0	20.0	30.8	30.8	8.0
DEX0448	027.nt.3	9622.0	14.3	14.3	20.0	20.0	30.8	30.8	8.0
DEX0448	027.nt.3	9623.0	14.3	14.3	20.0	20.0	30.8	30.8	8.0
DEX0448	027.nt.4	9622.0	14.3	14.3	20.0	20.0	30.8	30.8	8.0
DEX0448	027.nt.4	9623.0	14.3	14.3	20.0	20.0	30.8	30.8	8.0
DEX0448	027.nt.5	9622.0	14.3	14.3	20.0	20.0	30.8	30.8	8.0
DEX0448	027.nt.5	9623.0	14.3	14.3	20.0	20.0	30.8	30.8	8.0

DEX0448	029.nt.3	28734.0	3.6	3.6	40.0	40.0	15.4	15.4	12.0	12.0
DEX0448	029.nt.3	38381.0	0.0	0.0	10.0	11.1	7.7	7.7	0.0	0.0
DEX0448	029.nt.3	38382.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	029.nt.3	38383.0	0.0	0.0	10.0	10.0	0.0	0.0	4.0	4.0
DEX0448	029.nt.3	38384.0	0.0	0.0	10.0	10.0	0.0	0.0	4.0	4.0
DEX0448	029.nt.4	28733.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	029.nt.4	28734.0	3.6	3.6	40.0	40.0	15.4	15.4	12.0	12.0
DEX0448	029.nt.4	38381.0	0.0	0.0	10.0	11.1	7.7	7.7	0.0	0.0
DEX0448	029.nt.4	38382.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	029.nt.4	38383.0	0.0	0.0	10.0	10.0	0.0	0.0	4.0	4.0
DEX0448	029.nt.4	38384.0	0.0	0.0	10.0	10.0	0.0	0.0	4.0	4.0
DEX0448	029.nt.5	28733.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	029.nt.5	28734.0	3.6	3.6	40.0	40.0	15.4	15.4	12.0	12.0
DEX0448	029.nt.5	38381.0	0.0	0.0	10.0	11.1	7.7	7.7	0.0	0.0
DEX0448	029.nt.5	38382.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	029.nt.5	38383.0	0.0	0.0	10.0	10.0	0.0	0.0	4.0	4.0
DEX0448	029.nt.5	38384.0	0.0	0.0	10.0	10.0	0.0	0.0	4.0	4.0
DEX0448	029.nt.6	28733.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	029.nt.6	28734.0	3.6	3.6	40.0	40.0	15.4	15.4	12.0	12.0
DEX0448	029.nt.6	38381.0	0.0	0.0	10.0	11.1	7.7	7.7	0.0	0.0
DEX0448	029.nt.6	38382.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	029.nt.6	38383.0	0.0	0.0	10.0	10.0	0.0	0.0	4.0	4.0
DEX0448	029.nt.6	38384.0	0.0	0.0	10.0	10.0	0.0	0.0	4.0	4.0
DEX0448	029.nt.7	28733.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	029.nt.7	28734.0	3.6	3.6	40.0	40.0	15.4	15.4	12.0	12.0
DEX0448	029.nt.7	38381.0	0.0	0.0	10.0	11.1	7.7	7.7	0.0	0.0
DEX0448	029.nt.7	38382.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	029.nt.7	38383.0	0.0	0.0	10.0	10.0	0.0	0.0	4.0	4.0
DEX0448	029.nt.7	38384.0	0.0	0.0	10.0	10.0	0.0	0.0	4.0	4.0
DEX0448	030.nt.1	29603.0	3.6	3.6	10.0	10.0	0.0	0.0	8.0	8.0
DEX0448	030.nt.1	29604.0	7.1	7.4	10.0	10.0	0.0	0.0	12.0	12.5
DEX0448	030.nt.1	40867.0	7.1	8.0	20.0	20.0	15.4	16.7	8.0	8.7
DEX0448	030.nt.1	40868.0	7.1	7.1	30.0	30.0	15.4	15.4	12.0	12.0
DEX0448	032.nt.1	36179.0	17.9	18.5	10.0	11.1	7.7	8.3	20.0	20.8
DEX0448	032.nt.1	36180.0	10.7	11.1	10.0	11.1	7.7	8.3	12.0	12.5
DEX0448	032.nt.1	36182.0	3.6	4.2	0.0	0.0	7.7	10.0	0.0	0.0
DEX0448	032.nt.1	36184.0	39.3	39.3	30.0	30.0	46.2	46.2	32.0	32.0
DEX0448	032.nt.1	37227.0	7.1	8.3	0.0	0.0	7.7	11.1	4.0	4.5
DEX0448	032.nt.1	37228.0	3.6	4.0	0.0	0.0	7.7	10.0	0.0	0.0
DEX0448	032.nt.1	37239.0	28.6	28.6	30.0	30.0	38.5	38.5	24.0	24.0
DEX0448	032.nt.1	37240.0	32.1	33.3	40.0	40.0	46.2	46.2	28.0	29.2
DEX0448	035.nt.1	23862.0	14.3	14.3	30.0	30.0	23.1	23.1	16.0	16.0
DEX0448	035.nt.1	23863.0	25.0	25.0	30.0	30.0	30.8	30.8	24.0	24.0
DEX0448	035.nt.2	23862.0	14.3	14.3	30.0	30.0	23.1	23.1	16.0	16.0
DEX0448	035.nt.2	23863.0	25.0	25.0	30.0	30.0	30.8	30.8	24.0	24.0
DEX0448	035.nt.3	23862.0	14.3	14.3	30.0	30.0	23.1	23.1	16.0	16.0
DEX0448	035.nt.3	23863.0	25.0	25.0	30.0	30.0	30.8	30.8	24.0	24.0
DEX0448	035.nt.4	23862.0	14.3	14.3	30.0	30.0	23.1	23.1	16.0	16.0
DEX0448	035.nt.4	23863.0	25.0	25.0	30.0	30.0	30.8	30.8	24.0	24.0
DEX0448	036.nt.1	29532.0	17.9	17.9	50.0	50.0	38.5	38.5	20.0	20.0
DEX0448	036.nt.1	29533.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	036.nt.1	29534.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	036.nt.1	29540.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	036.nt.1	36811.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	036.nt.1	36812.0	0.0	0.0	10.0	12.5	7.7	11.1	0.0	0.0
DEX0448	036.nt.2	29532.0	17.9	17.9	50.0	50.0	38.5	38.5	20.0	20.0
DEX0448	036.nt.2	29539.0	0.0	0.0	10.0	11.1	7.7	7.7	0.0	0.0

DEX0448	036.nt.2	29540.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	036.nt.3	29532.0	17.9	17.9	50.0	50.0	38.5	38.5	20.0	20.0
DEX0448	036.nt.3	29539.0	0.0	0.0	10.0	11.1	7.7	7.7	0.0	0.0
DEX0448	036.nt.3	29540.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	037.nt.1	36341.0	7.1	7.1	20.0	20.0	30.8	30.8	0.0	0.0
DEX0448	037.nt.1	36342.0	3.6	3.8	20.0	20.0	23.1	25.0	0.0	0.0
DEX0448	037.nt.2	36341.0	7.1	7.1	20.0	20.0	30.8	30.8	0.0	0.0
DEX0448	037.nt.2	36342.0	3.6	3.8	20.0	20.0	23.1	25.0	0.0	0.0
DEX0448	037.nt.3	36341.0	7.1	7.1	20.0	20.0	30.8	30.8	0.0	0.0
DEX0448	037.nt.3	36342.0	3.6	3.8	20.0	20.0	23.1	25.0	0.0	0.0
DEX0448	037.nt.4	36341.0	7.1	7.1	20.0	20.0	30.8	30.8	0.0	0.0
DEX0448	037.nt.4	36342.0	3.6	3.8	20.0	20.0	23.1	25.0	0.0	0.0
DEX0448	037.nt.5	36341.0	7.1	7.1	20.0	20.0	30.8	30.8	0.0	0.0
DEX0448	037.nt.5	36342.0	3.6	3.8	20.0	20.0	23.1	25.0	0.0	0.0
DEX0448	037.nt.6	36341.0	7.1	7.1	20.0	20.0	30.8	30.8	0.0	0.0
DEX0448	037.nt.6	36342.0	3.6	3.8	20.0	20.0	23.1	25.0	0.0	0.0
DEX0448	038.nt.1	20895.0	28.6	28.6	30.0	33.3	84.6	91.7	0.0	0.0
DEX0448	038.nt.1	20896.0	28.6	28.6	40.0	40.0	92.3	92.3	0.0	0.0
DEX0448	039.nt.1	38855.0	28.6	28.6	50.0	50.0	23.1	23.1	40.0	40.0
DEX0448	040.nt.1	31347.0	32.1	33.3	30.0	30.0	53.8	53.8	20.0	20.8
DEX0448	040.nt.1	40755.0	28.6	28.6	30.0	30.0	53.8	53.8	16.0	16.0
DEX0448	040.nt.1	40756.0	32.1	32.1	30.0	30.0	53.8	53.8	20.0	20.0
DEX0448	040.nt.2	31346.0	35.7	35.7	40.0	40.0	61.5	61.5	24.0	24.0
DEX0448	040.nt.2	31347.0	32.1	33.3	30.0	30.0	53.8	53.8	20.0	20.8
DEX0448	040.nt.2	40755.0	28.6	28.6	30.0	30.0	53.8	53.8	16.0	16.0
DEX0448	040.nt.2	40756.0	32.1	32.1	30.0	30.0	53.8	53.8	20.0	20.0
DEX0448	040.nt.3	31346.0	35.7	35.7	40.0	40.0	61.5	61.5	24.0	24.0
DEX0448	040.nt.3	31347.0	32.1	33.3	30.0	30.0	53.8	53.8	20.0	20.8
DEX0448	040.nt.3	40755.0	28.6	28.6	30.0	30.0	53.8	53.8	16.0	16.0
DEX0448	040.nt.3	40756.0	32.1	32.1	30.0	30.0	53.8	53.8	20.0	20.0
DEX0448	040.nt.4	31346.0	35.7	35.7	40.0	40.0	61.5	61.5	24.0	24.0
DEX0448	040.nt.4	31347.0	32.1	33.3	30.0	30.0	53.8	53.8	20.0	20.8
DEX0448	040.nt.4	40755.0	28.6	28.6	30.0	30.0	53.8	53.8	16.0	16.0
DEX0448	040.nt.4	40756.0	32.1	32.1	30.0	30.0	53.8	53.8	20.0	20.0
DEX0448	040.nt.5	31346.0	35.7	35.7	40.0	40.0	61.5	61.5	24.0	24.0
DEX0448	040.nt.5	31347.0	32.1	33.3	30.0	30.0	53.8	53.8	20.0	20.8
DEX0448	040.nt.5	40755.0	28.6	28.6	30.0	30.0	53.8	53.8	16.0	16.0
DEX0448	040.nt.5	40756.0	32.1	32.1	30.0	30.0	53.8	53.8	20.0	20.0
DEX0448	040.nt.6	31346.0	35.7	35.7	40.0	40.0	61.5	61.5	24.0	24.0
DEX0448	040.nt.6	31347.0	32.1	33.3	30.0	30.0	53.8	53.8	20.0	20.8
DEX0448	040.nt.6	35943.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	040.nt.6	35944.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	040.nt.6	40755.0	28.6	28.6	30.0	30.0	53.8	53.8	16.0	16.0
DEX0448	040.nt.6	40756.0	32.1	32.1	30.0	30.0	53.8	53.8	20.0	20.0
DEX0448	040.nt.7	31346.0	35.7	35.7	40.0	40.0	61.5	61.5	24.0	24.0
DEX0448	040.nt.7	31347.0	32.1	33.3	30.0	30.0	53.8	53.8	20.0	20.8
DEX0448	040.nt.7	40755.0	28.6	28.6	30.0	30.0	53.8	53.8	16.0	16.0
DEX0448	040.nt.7	40756.0	32.1	32.1	30.0	30.0	53.8	53.8	20.0	20.0
DEX0448	041.nt.1	32216.0	50.0	50.0	60.0	60.0	61.5	61.5	48.0	48.0
DEX0448	042.nt.1	29231.0	17.9	18.5	10.0	10.0	23.1	23.1	12.0	12.5
DEX0448	042.nt.1	29232.0	17.9	17.9	10.0	10.0	15.4	15.4	16.0	16.0
DEX0448	042.nt.1	29271.0	21.4	21.4	30.0	30.0	30.8	30.8	20.0	20.0
DEX0448	042.nt.1	29272.0	21.4	21.4	40.0	40.0	30.8	30.8	24.0	24.0
DEX0448	042.nt.1	29290.0	32.1	32.1	30.0	30.0	46.2	46.2	24.0	24.0
DEX0448	043.nt.1	9043.0	14.3	14.3	30.0	30.0	23.1	23.1	16.0	16.0
DEX0448	044.nt.1	36803.0	32.1	32.1	20.0	20.0	0.0	0.0	44.0	44.0

Table 3.

DEX ID	Oligo Name	Cln Multi-Cancer ALL %up n=27	Cln Multi-Cancer ALL %valid up n=27	Cln Multi-Cancer ASC %up n=14	Cln Multi-Cancer ASC %valid up n=14	Cln Multi-Cancer RS %up n=13	Cln Multi-Cancer RS %valid up n=13
DEX0448 016.nt.1	5354.0	11.1	17.6	0.0	0.0	23.1	42.9
DEX0448 019.nt.1	1045.0	25.9	25.9	14.3	14.3	38.5	38.5
DEX0448 026.nt.1	78479.0	18.5	20.8	35.7	41.7	0.0	0.0
DEX0448 026.nt.1	78479.1	14.8	15.4	28.6	28.6	0.0	0.0
DEX0448 026.nt.2	78479.0	18.5	20.8	35.7	41.7	0.0	0.0
DEX0448 026.nt.2	78479.1	14.8	15.4	28.6	28.6	0.0	0.0
DEX0448 028.nt.1	5305.0	14.8	14.8	28.6	28.6	0.0	0.0
DEX0448 028.nt.1	5306.0	14.8	14.8	28.6	28.6	0.0	0.0
DEX0448 028.nt.2	5305.0	14.8	14.8	28.6	28.6	0.0	0.0
DEX0448 028.nt.2	5306.0	14.8	14.8	28.6	28.6	0.0	0.0
DEX0448 034.nt.1	5354.0	11.1	17.6	0.0	0.0	23.1	42.9
DEX0448 044.nt.1	42013.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 044.nt.1	42013.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 044.nt.1	42013.2	0.0	0.0	0.0	0.0	0.0	0.0

BREAST CANCER CHIPS

For breast cancer two different chip designs were evaluated with overlapping sets of a total of 36 samples, comparing the expression patterns of breast cancer derived polyA⁺ RNA to polyA⁺ RNA isolated from a pool of 10 normal breast tissues. For the Breast Array Chip, all 36 samples (9 stage I cancers, 23 stage II cancers, 4 stage III cancers) were analyzed. These samples also represented 10 Grade 1/2 and 26 Grade 3 cancers. The histopathologic grades for cancer are classified as follows: GX, cannot be assessed; G1, well differentiated; G2, moderately differentiated; G3, poorly differentiated; and G4, undifferentiated. AJCC Cancer Staging Handbook, pp. 9, (5th Ed, 1998). Samples were further grouped based on the expression patterns of the known breast cancer associated genes Her2 and ER α (10 HER2 up, 26 HER2 not up, 20 ER up and 16 ER not up) and for the Multi-Cancer Array Chip, a subset of 20 of these samples (9 stage I cancers, 8 stage II cancers, 3 stage III cancers) were assessed.

The results for the statistically significant up-regulated genes on the Breast Array Chip are shown in Tables 4 and 5. The results for the statistically significant up-regulated genes on the Multi-Cancer Array Chip are shown in Table 6. The first two columns of each table contain information about the sequence itself (Seq ID, Oligo Name), the next columns show the results obtained for all ("ALL") breast cancer samples, cancers corresponding to stage I ("ST1"), stages II and III ("ST2,3"), grades 1 and 2 ("GR1,2"), grade 3 ("GR3"), cancers exhibiting up-regulation of Her2 ("HER2up")

- or ER α ("ERup") or those not exhibiting up-regulation of Her2 ("NOT HER2up") or ER α ("NOT ERup"). '%up' indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed (n=36 for Colon Array Chip, n=20 for the Multi-Cancer Array Chip), '%valid up' indicates the percentage of experiments with
- 5 valid expression values in which up-regulation of at least 2-fold was observed.

Table 4.

DEX ID	Oligo Name	Mam ALL %up n=36	Mam ALL %valid up n=36	Mam ST1 %up n=9	Mam ST1 %valid up n=9	Mam ST2,3 %up n=27	Mam ST2,3 %valid up n=27	Mam GR1,2 %up n=10	Mam GR1,2 %valid up n=10	Mam GR3 %up n=26	Mam GR3 %valid up n=26
DEX0448 010.nt.1	17869.0	8.3	8.3	11.1	11.1	7.4	7.4	0.0	0.0	11.5	11.5
DEX0448 010.nt.1	31466.0	8.3	8.3	0.0	0.0	11.1	11.1	0.0	0.0	11.5	11.5
DEX0448 011.nt.1	20711.0	16.7	16.7	0.0	0.0	22.2	22.2	0.0	0.0	23.1	23.1
DEX0448 016.nt.1	22545.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.1	22546.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.1	26543.0	2.8	2.8	0.0	0.0	3.7	3.7	0.0	0.0	3.8	3.8
DEX0448 016.nt.2	22545.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.2	22546.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.2	26543.0	2.8	2.8	0.0	0.0	3.7	3.7	0.0	0.0	3.8	3.8
DEX0448 016.nt.3	22545.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.3	22546.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.3	26543.0	2.8	2.8	0.0	0.0	3.7	3.7	0.0	0.0	3.8	3.8
DEX0448 017.nt.1	29052.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 018.nt.1	13245.0	11.1	11.1	22.2	22.2	7.4	7.4	10.0	10.0	11.5	11.5
DEX0448 018.nt.1	13299.0	11.1	11.1	11.1	11.1	11.1	11.1	0.0	0.0	15.4	15.4
DEX0448 018.nt.1	13320.0	5.6	5.6	11.1	11.1	3.7	3.7	0.0	0.0	7.7	7.7
DEX0448 018.nt.1	13321.0	11.1	11.1	11.1	11.1	11.1	11.1	0.0	0.0	15.4	15.4
DEX0448 018.nt.1	13323.0	13.9	13.9	11.1	11.1	14.8	14.8	0.0	0.0	19.2	19.2
DEX0448 018.nt.1	13337.0	11.1	11.1	22.2	22.2	7.4	7.4	10.0	10.0	11.5	11.5
DEX0448 018.nt.1	13338.0	5.6	5.7	11.1	11.1	3.7	3.8	0.0	0.0	7.7	8.0
DEX0448 022.nt.1	32150.0	36.1	36.1	55.6	55.6	29.6	29.6	20.0	20.0	42.3	42.3
DEX0448 022.nt.1	32151.0	22.2	22.2	44.4	44.4	14.8	14.8	10.0	10.0	26.9	26.9
DEX0448 022.nt.2	32150.0	36.1	36.1	55.6	55.6	29.6	29.6	20.0	20.0	42.3	42.3
DEX0448 022.nt.2	32151.0	22.2	22.2	44.4	44.4	14.8	14.8	10.0	10.0	26.9	26.9
DEX0448 027.nt.1	15310.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.1	15311.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.1	15833.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.1	15834.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.1	15835.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.1	15836.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.1	15845.0	2.8	2.9	0.0	0.0	3.7	3.7	0.0	0.0	3.8	4.0
DEX0448 027.nt.1	15846.0	2.8	2.8	0.0	0.0	3.7	3.7	0.0	0.0	3.8	3.8
DEX0448 027.nt.1	32136.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.1	32137.0	2.8	2.8	0.0	0.0	3.7	3.7	0.0	0.0	3.8	3.8
DEX0448 027.nt.2	15310.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.2	15311.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.2	15833.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.2	15834.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.2	15835.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.2	15836.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.2	15845.0	2.8	2.9	0.0	0.0	3.7	3.7	0.0	0.0	3.8	4.0

DEX0448 027.nt.2	15846.0	2.8	2.8	0.0	0.0	3.7	3.7	0.0	0.0	3.8	3.8
DEX0448 027.nt.2	32136.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.2	32137.0	2.8	2.8	0.0	0.0	3.7	3.7	0.0	0.0	3.8	3.8
DEX0448 027.nt.3	15310.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.3	15311.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.3	15833.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.3	15834.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.3	15835.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.3	15836.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.3	15845.0	2.8	2.9	0.0	0.0	3.7	3.7	0.0	0.0	3.8	4.0
DEX0448 027.nt.3	15846.0	2.8	2.8	0.0	0.0	3.7	3.7	0.0	0.0	3.8	3.8
DEX0448 027.nt.3	32136.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.3	32137.0	2.8	2.8	0.0	0.0	3.7	3.7	0.0	0.0	3.8	3.8
DEX0448 027.nt.4	15310.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.4	15311.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.4	15833.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.4	15834.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.4	15835.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.4	15836.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.4	32136.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.4	32137.0	2.8	2.8	0.0	0.0	3.7	3.7	0.0	0.0	3.8	3.8
DEX0448 027.nt.5	15310.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.5	15311.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.5	15833.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.5	15834.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.5	15835.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.5	15836.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.5	32136.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.5	32137.0	2.8	2.8	0.0	0.0	3.7	3.7	0.0	0.0	3.8	3.8
DEX0448 027.nt.6	15310.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.6	15311.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.6	15835.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.6	15836.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.6	15845.0	2.8	2.9	0.0	0.0	3.7	3.7	0.0	0.0	3.8	4.0
DEX0448 027.nt.6	15846.0	2.8	2.8	0.0	0.0	3.7	3.7	0.0	0.0	3.8	3.8
DEX0448 027.nt.6	32136.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 027.nt.6	32137.0	2.8	2.8	0.0	0.0	3.7	3.7	0.0	0.0	3.8	3.8
DEX0448 033.nt.1	22294.0	2.8	2.8	11.1	11.1	0.0	0.0	10.0	10.0	0.0	0.0
DEX0448 034.nt.1	22545.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 034.nt.1	22546.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 035.nt.1	40309.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 043.nt.1	22586.0	8.3	8.3	0.0	0.0	11.1	11.1	0.0	0.0	11.5	11.5
DEX0448 043.nt.2	22585.0	8.3	8.3	0.0	0.0	11.1	11.1	0.0	0.0	11.5	11.5
DEX0448 043.nt.2	22586.0	8.3	8.3	0.0	0.0	11.1	11.1	0.0	0.0	11.5	11.5

Table 5.

DEX ID	Oligo Name	Mam HER2 up %up n=10	Mam HER2 up %valid up n=10	Mam NOT HER2 up %up n=26	Mam NOT HER2 up %valid up n=26	Mam ER up %up n=20	Mam ER up %valid up n=20	Mam NOT ER up %up n=16	Mam NOT ER up %valid up n=16
DEX0448 010.nt.1	17869.0	0.0	0.0	11.5	11.5	0.0	0.0	18.8	18.8
DEX0448 010.nt.1	31466.0	0.0	0.0	11.5	11.5	0.0	0.0	18.8	18.8
DEX0448 011.nt.1	20711.0	20.0	20.0	15.4	15.4	25.0	25.0	6.2	6.2

[illegible]

DEX0448	027.nt.4	32136.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	027.nt.4	32137.0	10.0	10.0	0.0	0.0	0.0	0.0	0.0	6.2	6.2
DEX0448	027.nt.5	15310.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	027.nt.5	15311.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	027.nt.5	15833.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	027.nt.5	15834.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	027.nt.5	15835.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	027.nt.5	15836.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	027.nt.5	32136.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	027.nt.5	32137.0	10.0	10.0	0.0	0.0	0.0	0.0	0.0	6.2	6.2
DEX0448	027.nt.6	15310.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	027.nt.6	15311.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	027.nt.6	15835.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	027.nt.6	15836.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	027.nt.6	15845.0	10.0	10.0	0.0	0.0	0.0	0.0	0.0	6.2	6.2
DEX0448	027.nt.6	15846.0	10.0	10.0	0.0	0.0	0.0	0.0	0.0	6.2	6.2
DEX0448	027.nt.6	32136.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	027.nt.6	32137.0	10.0	10.0	0.0	0.0	0.0	0.0	0.0	6.2	6.2
DEX0448	033.nt.1	22294.0	0.0	0.0	3.8	3.8	5.0	5.0	0.0	0.0	0.0
DEX0448	034.nt.1	22545.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	034.nt.1	22546.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	035.nt.1	40309.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448	043.nt.1	22586.0	0.0	0.0	11.5	11.5	5.0	5.0	12.5	12.5	12.5
DEX0448	043.nt.2	22585.0	0.0	0.0	11.5	11.5	5.0	5.0	12.5	12.5	12.5
DEX0448	043.nt.2	22586.0	0.0	0.0	11.5	11.5	5.0	5.0	12.5	12.5	12.5

Table 6.

DEX ID	Oligo Name	Mam Multi-Cancer ALL %up n=20	Mam Multi-Cancer ALL %valid up n=20	Mam Multi-Cancer ST1 %up n=9	Mam Multi-Cancer ST1 %valid up n=9	Mam Multi-Cancer ST2,3 %up n=11	Mam Multi-Cancer ST2,3 %valid up n=11
DEX0448	016.nt.1	5354.0	0.0	0.0	0.0	0.0	0.0
DEX0448	019.nt.1	1045.0	0.0	0.0	0.0	0.0	0.0
DEX0448	026.nt.1	78479.0	0.0	0.0	0.0	0.0	0.0
DEX0448	026.nt.1	78479.1	0.0	0.0	0.0	0.0	0.0
DEX0448	026.nt.2	78479.0	0.0	0.0	0.0	0.0	0.0
DEX0448	026.nt.2	78479.1	0.0	0.0	0.0	0.0	0.0
DEX0448	028.nt.1	5305.0	0.0	0.0	0.0	0.0	0.0
DEX0448	028.nt.1	5306.0	0.0	0.0	0.0	0.0	0.0
DEX0448	028.nt.2	5305.0	0.0	0.0	0.0	0.0	0.0
DEX0448	028.nt.2	5306.0	0.0	0.0	0.0	0.0	0.0
DEX0448	034.nt.1	5354.0	0.0	0.0	0.0	0.0	0.0
DEX0448	044.nt.1	42013.0	0.0	0.0	0.0	0.0	0.0
DEX0448	044.nt.1	42013.1	0.0	0.0	0.0	0.0	0.0
DEX0448	044.nt.1	42013.2	0.0	0.0	0.0	0.0	0.0

LUNG CANCER CHIPS

- 5 For lung cancer two different chip designs were evaluated with overlapping sets of a total of 29 samples, comparing the expression patterns of lung cancer derived polyA+ RNA to polyA+ RNA isolated from a pool of 12 normal lung tissues. For the Lung Array

Chip all 29 samples (15 squamous cell carcinomas and 14 adenocarcinomas including 14 stage I and 15 stage II/III cancers) were analyzed and for the Multi-Cancer Array Chip a subset of 22 of these samples (10 squamous cell carcinomas, 12 adenocarcinomas) were assessed.

5 The results for the statistically significant up-regulated genes on the Lung Array
Chip are shown in Table 7. The results for the statistically significant up-regulated genes
on the Multi-Cancer Array Chip are shown in Table 8. The first two columns of each
table contain information about the sequence itself (DEX ID, Oligo Name), the next
columns show the results obtained for all ("ALL") lung cancer samples, squamous cell
10 carcinomas ("SQ"), adenocarcinomas ("AD"), or cancers corresponding to stage I
("ST1"), or stages II and III ("ST2,3"). '%up' indicates the percentage of all experiments
in which up-regulation of at least 2-fold was observed (n=29 for Lung Array Chip, n=22
for Multi-Cancer Array Chip), '%valid up' indicates the percentage of experiments with
valid expression values in which up-regulation of at least 2-fold was observed.

Table 7.

[illegible]

DEX0448 028.nt.2	1264.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 028.nt.2	1265.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 028.nt.2	1320.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 034.nt.1	5353.0	13.8	13.8	20.0	20.0	7.1	7.1	14.3	14.3	13.3	13.3
DEX0448 035.nt.1	3729.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 040.nt.2	7434.0	10.3	11.5	6.7	7.1	14.3	16.7	14.3	16.7	6.7	7.1
DEX0448 040.nt.2	7435.0	10.3	12.5	6.7	8.3	14.3	16.7	14.3	15.4	6.7	9.1
DEX0448 040.nt.3	7434.0	10.3	11.5	6.7	7.1	14.3	16.7	14.3	16.7	6.7	7.1
DEX0448 040.nt.3	7435.0	10.3	12.5	6.7	8.3	14.3	16.7	14.3	15.4	6.7	9.1
DEX0448 040.nt.4	7434.0	10.3	11.5	6.7	7.1	14.3	16.7	14.3	16.7	6.7	7.1
DEX0448 040.nt.4	7435.0	10.3	12.5	6.7	8.3	14.3	16.7	14.3	15.4	6.7	9.1
DEX0448 040.nt.5	7434.0	10.3	11.5	6.7	7.1	14.3	16.7	14.3	16.7	6.7	7.1
DEX0448 040.nt.5	7435.0	10.3	12.5	6.7	8.3	14.3	16.7	14.3	15.4	6.7	9.1
DEX0448 040.nt.6	7434.0	10.3	11.5	6.7	7.1	14.3	16.7	14.3	16.7	6.7	7.1
DEX0448 040.nt.6	7435.0	10.3	12.5	6.7	8.3	14.3	16.7	14.3	15.4	6.7	9.1
DEX0448 042.nt.1	7412.0	13.8	13.8	20.0	20.0	7.1	7.1	14.3	14.3	13.3	13.3
DEX0448 042.nt.1	7413.0	17.2	17.2	20.0	20.0	14.3	14.3	21.4	21.4	13.3	13.3
DEX0448 043.nt.2	939.0	3.4	4.0	0.0	0.0	7.1	7.7	0.0	0.0	6.7	6.7
DEX0448 043.nt.2	940.0	3.4	3.4	0.0	0.0	7.1	7.1	0.0	0.0	6.7	6.7
DEX0448 043.nt.2	943.0	3.4	3.4	0.0	0.0	7.1	7.1	0.0	0.0	6.7	6.7
DEX0448 043.nt.2	944.0	3.4	3.4	0.0	0.0	7.1	7.1	0.0	0.0	6.7	6.7
DEX0448 043.nt.2	946.0	3.4	3.4	0.0	0.0	7.1	7.1	0.0	0.0	6.7	6.7
DEX0448 043.nt.2	960.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 043.nt.2	966.0	3.4	3.4	0.0	0.0	7.1	7.1	0.0	0.0	6.7	6.7
DEX0448 044.nt.1	42013.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 8.

DEX ID	Oligo Name	Lng Multi-Cancer ALL %up n=22	Lng Multi-Cancer ALL %valid up n=22	Lng Multi-Cancer SQ %up n=10	Lng Multi-Cancer SQ %valid up n=10	Lng Multi-Cancer AD %up n=12	Lng Multi-Cancer AD %valid up n=12
DEX0448 016.nt.1	5354.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 019.nt.1	1045.0	4.5	4.5	0.0	0.0	8.3	8.3
DEX0448 026.nt.1	78479.0	31.8	36.8	10.0	12.5	50.0	54.5
DEX0448 026.nt.1	78479.1	36.4	36.4	10.0	10.0	58.3	58.3
DEX0448 026.nt.2	78479.0	31.8	36.8	10.0	12.5	50.0	54.5
DEX0448 026.nt.2	78479.1	36.4	36.4	10.0	10.0	58.3	58.3
DEX0448 028.nt.1	5305.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 028.nt.1	5306.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 028.nt.2	5305.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 028.nt.2	5306.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 034.nt.1	5354.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 044.nt.1	42013.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 044.nt.1	42013.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 044.nt.1	42013.2	0.0	0.0	0.0	0.0	0.0	0.0

OVARIAN CANCER CHIPS

5

For ovarian cancer two different chip designs were evaluated with overlapping sets of a total of 19 samples, comparing the expression patterns of ovarian cancer derived total RNA to total RNA isolated from a pool of 9 normal ovarian tissues. For the Multi-Cancer Array Chip, all 19 samples (14 invasive carcinomas, 5 low malignant potential samples

were analyzed and for the Ovarian Array Chip, a subset of 17 of these samples (13 invasive carcinomas, 4 low malignant potential samples) were assessed.

The results for the statistically significant up-regulated genes on the Ovarian Array Chip are shown in Table 9. The results for the Multi-Cancer Array Chip are shown in Table 10. The first two columns of each table contain information about the sequence itself (DEX ID, Oligo Name), the next columns show the results obtained for all ("ALL") ovarian cancer samples, invasive carcinomas ("INV") and low malignant potential ("LMP") samples. '%up' indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed (n=19 for the Multi-Cancer Array Chip, n=17 for the Ovarian Array Chip), '%valid up' indicates the percentage of experiments with valid expression values in which up-regulation of at least 2-fold was observed.

Table 9.

DEX ID	Oligo Name	Ovr ALL %up n=17	Ovr ALL %valid up n=17	Ovr INV %up n=13	Ovr INV %valid up n=13	Ovr LMP %up n=4	Ovr LMP %valid up n=4
DEX0448 010.nt.1	9880.01	17.6	21.4	15.4	18.2	25.0	33.3
DEX0448 010.nt.1	9880.02	17.6	18.8	15.4	16.7	25.0	25.0
DEX0448 011.nt.1	22483.01	11.8	11.8	7.7	7.7	25.0	25.0
DEX0448 011.nt.1	22483.02	11.8	11.8	7.7	7.7	25.0	25.0
DEX0448 016.nt.1	21617.01	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.1	21617.02	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.1	21619.01	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.1	21619.02	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.2	21613.01	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.2	21613.02	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.2	21617.01	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.2	21617.02	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.2	21619.01	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.2	21619.02	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.3	21617.01	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.3	21617.02	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.3	21619.01	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 016.nt.3	21619.02	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 034.nt.1	21613.01	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 034.nt.1	21613.02	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 034.nt.1	21619.01	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 034.nt.1	21619.02	0.0	0.0	0.0	0.0	0.0	0.0

15 Table 10.

DEX ID	Oligo Name	Ovr Multi-Cancer ALL %up n=19	Ovr Multi-Cancer ALL %valid up n=19	Ovr Multi-Cancer INV %up n=14	Ovr Multi-Cancer INV %valid up n=14	Ovr Multi-Cancer LMP %up n=5	Ovr Multi-Cancer LMP %valid up n=5
DEX0448 016.nt.1	5354.0	0.0	0.0	0.0	0.0	0.0	0.0

DEX0448 019.nt.1	1045.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 026.nt.1	78479.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 026.nt.1	78479.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 026.nt.2	78479.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 026.nt.2	78479.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 028.nt.1	5305.0	5.3	5.6	7.1	7.7	0.0	0.0
DEX0448 028.nt.1	5306.0	5.3	5.6	7.1	7.1	0.0	0.0
DEX0448 028.nt.2	5305.0	5.3	5.6	7.1	7.7	0.0	0.0
DEX0448 028.nt.2	5306.0	5.3	5.6	7.1	7.1	0.0	0.0
DEX0448 034.nt.1	5354.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 044.nt.1	42013.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 044.nt.1	42013.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0448 044.nt.1	42013.2	0.0	0.0	0.0	0.0	0.0	0.0

PROSTATE CANCER

For prostate cancer three different chip designs were evaluated with overlapping sets of a total of 29 samples, comparing the expression patterns of prostate cancer or benign disease derived total RNA to total RNA isolated from a pool of 35 normal prostate tissues. For the Prostate 1 Array and Prostate 2 Array Chips all 29 samples (17 prostate cancer samples 12 non-malignant disease samples) were analyzed. For the Multi-Cancer Array Chip a subset of 28 of these samples (16 prostate cancer samples, 12 non-malignant disease samples) were analyzed.

The results for the statistically significant up-regulated genes on the Prostate1 Array Chip and Prostate2 Array Chip are shown in Table 11. The results for the statistically significant up-regulated genes on the Multi-Cancer Array Chip are shown in Table 12. The first two columns of each table contain information about the sequence itself (DEX ID, Oligo Name), the next columns show the results obtained for prostate cancer samples ("CAN") or non-malignant disease samples ("DIS"). '%up' indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed (n=29 for the Prostate2 Array Chip and the Multi-Cancer Array Chip), '%valid up' indicates the percentage of experiments with valid expression values in which up-regulation of at least 2-fold was observed.

Table 11.

DEX ID	Oligo Name	Pro CAN %up n=17	Pro CAN %valid up n=17	Pro DIS %up n=12	Pro2 DIS %valid up n=12
DEX0448 010.nt.1	38723.02	0.0	0.0	0.0	0.0
DEX0448 010.nt.1	38723.03	0.0	0.0	0.0	0.0
DEX0448 014.nt.1	26889.01	0.0	0.0	0.0	0.0
DEX0448 014.nt.1	26889.02	0.0	0.0	0.0	0.0
DEX0448 014.nt.1	27099.01	0.0	0.0	0.0	0.0

200

DEX0448 014.nt.1	27099.02	0.0	0.0	0.0	0.0
DEX0448 018.nt.1	32638.01	0.0	0.0	0.0	0.0
DEX0448 018.nt.1	32638.02	0.0	0.0	0.0	0.0
DEX0448 018.nt.1	32638.03	0.0	0.0	8.3	11.1
DEX0448 022.nt.1	29143.01	11.8	11.8	8.3	8.3
DEX0448 022.nt.1	29143.02	5.9	5.9	8.3	9.1
DEX0448 022.nt.1	29143.03	5.9	5.9	8.3	8.3
DEX0448 022.nt.2	29143.01	11.8	11.8	8.3	8.3
DEX0448 022.nt.2	29143.02	5.9	5.9	8.3	9.1
DEX0448 022.nt.2	29143.03	5.9	5.9	8.3	8.3
DEX0448 041.nt.1	28451.01	0.0	0.0	0.0	0.0
DEX0448 041.nt.1	28451.02	0.0	0.0	0.0	0.0
DEX0448 041.nt.1	32716.01	0.0	0.0	0.0	0.0
DEX0448 041.nt.1	32716.02	0.0	0.0	0.0	0.0
DEX0448 041.nt.1	32716.03	0.0	0.0	8.3	14.3
DEX0448 041.nt.1	32718.01	0.0	0.0	0.0	0.0
DEX0448 041.nt.1	32718.02	0.0	0.0	0.0	0.0
DEX0448 041.nt.1	32718.03	0.0	0.0	0.0	0.0
DEX0448 041.nt.1	32720.01	0.0	0.0	0.0	0.0
DEX0448 041.nt.1	32720.02	0.0	0.0	0.0	0.0
DEX0448 041.nt.1	32720.03	0.0	0.0	0.0	0.0
DEX0448 041.nt.1	32724.01	0.0	0.0	0.0	0.0
DEX0448 041.nt.1	32724.02	0.0	0.0	0.0	0.0
DEX0448 041.nt.1	32724.03	0.0	0.0	0.0	0.0
DEX0448 041.nt.2	32716.01	0.0	0.0	0.0	0.0
DEX0448 041.nt.2	32716.02	0.0	0.0	0.0	0.0
DEX0448 041.nt.2	32716.03	0.0	0.0	8.3	14.3

Table 12.

DEX ID	Oligo Name	Pro Multi-Cancer CAN %up n=16	Pro Multi-Cancer CAN %valid up n=16	Pro Multi-Cancer DIS %up n=12	Pro Multi-Cancer DIS %valid up n=12
DEX0448 016.nt.1	5354.0	0.0	0.0	0.0	0.0
DEX0448 019.nt.1	1045.0	0.0	0.0	0.0	0.0
DEX0448 026.nt.1	78479.0	0.0	0.0	0.0	0.0
DEX0448 026.nt.1	78479.1	0.0	0.0	0.0	0.0
DEX0448 026.nt.2	78479.0	0.0	0.0	0.0	0.0
DEX0448 026.nt.2	78479.1	0.0	0.0	0.0	0.0
DEX0448 028.nt.1	5305.0	0.0	0.0	0.0	0.0
DEX0448 028.nt.1	5306.0	0.0	0.0	0.0	0.0
DEX0448 028.nt.2	5305.0	0.0	0.0	0.0	0.0
DEX0448 028.nt.2	5306.0	0.0	0.0	0.0	0.0
DEX0448 034.nt.1	5354.0	0.0	0.0	0.0	0.0
DEX0448 044.nt.1	42013.0	0.0	0.0	0.0	0.0
DEX0448 044.nt.1	42013.1	0.0	0.0	0.0	0.0
DEX0448 044.nt.1	42013.2	0.0	0.0	0.0	0.0

- SEQ ID NO: 1-95 was up-regulated on various tissue microarrays. Accordingly,
- 5 nucleotide SEQ ID NO: 1-95 or the encoded protein SEQ ID NO: 96-237 may be used as a cancer therapeutic or diagnostic target for the tissues in which expression is shown.

The following table lists the location (Oligo Location) where the microarray oligos (Oligo ID) map on the transcripts (DEX ID) of the present invention. Each Oligo ID may have been printed multiple times on a single chip as replicates. The Oligo Name is an exemplary replicate (e.g. 1000.01) for the Oligo ID (e.g. 1000), and data from other replicates (e.g. 1000.02, 1000.03) may be reported. Additionally, the Array (Chip Name) that each oligo and oligo replicates were printed on is included.

DEX ID	Oligo ID	Oligo Name	Chip Name	Oligo Location
DEX0448_001.nt.1	34940	34940.0	Colon array	404-463
DEX0448_002.nt.1	39957	39957.0	Colon array	172-231
DEX0448_002.nt.1	39958	39958.0	Colon array	124-183
DEX0448_003.nt.1	32057	32057.0	Colon array	519-578
DEX0448_004.nt.1	41210	41210.0	Colon array	424-483
DEX0448_005.nt.2	36243	36243.0	Colon array	1134-1193
DEX0448_006.nt.1	8410	8410.0	Colon array	110-169
DEX0448_007.nt.1	26449	26449.0	Colon array	431-490
DEX0448_008.nt.1	32851	32851.0	Colon array	209-268
DEX0448_009.nt.1	40809	40809.0	Colon array	1893-1952
DEX0448_009.nt.1	34680	34680.0	Colon array	834-893
DEX0448_009.nt.1	40810	40810.0	Colon array	1853-1912
DEX0448_010.nt.1	28423	28423.0	Colon array	49-108
DEX0448_010.nt.1	953	953.0	Lung array	243-302
DEX0448_010.nt.1	956	956.0	Multi-Cancer array	45-104
DEX0448_010.nt.1	17869	17869.0	Breast array	44-103
DEX0448_010.nt.1	955	955.0	Lung array	50-109
DEX0448_010.nt.1	9880	9880.01	Ovarian array	49-108
DEX0448_010.nt.1	38723	38723.01	Prostate2 array	50-109
DEX0448_010.nt.1	835	835.0	Lung array	403-462
DEX0448_010.nt.1	5749	5749.0	Lung array	1038-1097
DEX0448_010.nt.1	31466	31466.0	Breast array	413-472
DEX0448_010.nt.1	836	836.0	Lung array	383-442
DEX0448_011.nt.1	22483	22483.02	Ovarian array	164-223
DEX0448_011.nt.1	20711	20711.0	Breast array	164-223
DEX0448_011.nt.1	8312	8312.0	Colon array	245-304
DEX0448_013.nt.1	16006	16006.0	Colon array	413-472
DEX0448_014.nt.1	27099	27099.01	Prostate1 array	257-316
DEX0448_014.nt.1	37378	37378.0	Colon array	137-196
DEX0448_014.nt.1	37376	37376.0	Colon array	137-196
DEX0448_014.nt.1	26889	26889.02	Prostate1 array	170-229
DEX0448_015.nt.1	38996	38996.0	Colon array	907-966
DEX0448_015.nt.1	33348	33348.0	Colon array	476-535
DEX0448_016.nt.1	22545	22545.0	Breast array	444-503
DEX0448_016.nt.1	26543	26543.0	Breast array	202-261
DEX0448_016.nt.1	21617	21617.01	Ovarian array	207-266
DEX0448_016.nt.1	5354	5354.0	Multi-Cancer array	434-493

DEX0448_016.nt.1	22546	22546.0	Breast array	414-473
DEX0448_016.nt.1	21619	21619.01	Ovarian array	295-354
DEX0448_016.nt.1	5353	5353.0	Lung array	445-504
DEX0448_016.nt.2	26543	26543.0	Breast array	439-498
DEX0448_016.nt.2	22545	22545.0	Breast array	681-740
DEX0448_016.nt.2	21619	21619.01	Ovarian array	532-591
DEX0448_016.nt.2	21613	21613.02	Ovarian array	815-874
DEX0448_016.nt.2	22546	22546.0	Breast array	651-710
DEX0448_016.nt.2	5354	5354.0	Multi-Cancer array	671-730
DEX0448_016.nt.2	21617	21617.01	Ovarian array	444-503
DEX0448_016.nt.2	5353	5353.0	Lung array	682-741
DEX0448_016.nt.3	26543	26543.0	Breast array	68-127
DEX0448_016.nt.3	22545	22545.0	Breast array	310-369
DEX0448_016.nt.3	21617	21617.01	Ovarian array	73-132
DEX0448_016.nt.3	5353	5353.0	Lung array	311-370
DEX0448_016.nt.3	22546	22546.0	Breast array	280-339
DEX0448_016.nt.3	5354	5354.0	Multi-Cancer array	300-359
DEX0448_016.nt.3	21619	21619.01	Ovarian array	161-220
DEX0448_017.nt.1	37895	37895.0	Colon array	863-922
DEX0448_017.nt.1	37227	37227.0	Colon array	724-783
DEX0448_017.nt.1	37239	37239.0	Colon array	319-378
DEX0448_017.nt.1	29052	29052.0	Breast array	871-930
DEX0448_017.nt.1	36180	36180.0	Colon array	551-610
DEX0448_017.nt.1	36184	36184.0	Colon array	218-277
DEX0448_017.nt.1	36182	36182.0	Colon array	724-783
DEX0448_017.nt.1	36179	36179.0	Colon array	592-651
DEX0448_017.nt.1	37228	37228.0	Colon array	694-753
DEX0448_017.nt.1	36181	36181.0	Colon array	852-911
DEX0448_017.nt.1	37240	37240.0	Colon array	299-358
DEX0448_018.nt.1	13323	13323.0	Breast array	902-961
DEX0448_018.nt.1	13245	13245.0	Breast array	1209-1268
DEX0448_018.nt.1	13337	13337.0	Breast array	1195-1254
DEX0448_018.nt.1	13299	13299.0	Breast array	853-912
DEX0448_018.nt.1	13338	13338.0	Breast array	1165-1224
DEX0448_018.nt.1	41823	41823.0	Colon array	544-603
DEX0448_018.nt.1	41824	41824.0	Colon array	495-554
DEX0448_018.nt.1	13320	13320.0	Breast array	1147-1206
DEX0448_018.nt.1	32638	32638.03	Prostate2 array	897-956
DEX0448_018.nt.1	13321	13321.0	Breast array	885-944
DEX0448_019.nt.1	1045	1045.0	Multi-Cancer array	144-203
DEX0448_019.nt.1	1044	1044.0	Lung array	154-213
DEX0448_020.nt.1	39656	39656.0	Colon array	1033-1092
DEX0448_020.nt.1	39655	39655.0	Colon array	1201-1260
DEX0448_021.nt.1	35218	35218.0	Colon array	1621-1680
DEX0448_021.nt.1	31155	31155.0	Colon array	892-951
DEX0448_021.nt.1	30870	30870.0	Colon array	1621-1680
DEX0448_021.nt.1	31146	31146.0	Colon array	603-662

DEX0448_021.nt.1	30819	30819.0	Colon array	1621-1680
DEX0448_021.nt.1	31157	31157.0	Colon array	1611-1670
DEX0448_021.nt.1	31154	31154.0	Colon array	1364-1423
DEX0448_021.nt.1	30931	30931.0	Colon array	653-712
DEX0448_021.nt.2	30870	30870.0	Colon array	1926-1985
DEX0448_021.nt.2	31154	31154.0	Colon array	1669-1728
DEX0448_021.nt.2	31157	31157.0	Colon array	1916-1975
DEX0448_021.nt.2	30819	30819.0	Colon array	1926-1985
DEX0448_021.nt.2	31146	31146.0	Colon array	603-662
DEX0448_021.nt.2	30931	30931.0	Colon array	653-712
DEX0448_021.nt.2	31155	31155.0	Colon array	892-951
DEX0448_021.nt.2	35218	35218.0	Colon array	1926-1985
DEX0448_021.nt.3	31154	31154.0	Colon array	207-266
DEX0448_021.nt.3	35218	35218.0	Colon array	464-523
DEX0448_021.nt.3	30819	30819.0	Colon array	464-523
DEX0448_021.nt.3	31157	31157.0	Colon array	454-513
DEX0448_021.nt.3	30870	30870.0	Colon array	464-523
DEX0448_022.nt.1	37415	37415.0	Colon array	452-511
DEX0448_022.nt.1	792	792.0	Lung array	452-511
DEX0448_022.nt.1	36878	36878.0	Colon array	502-561
DEX0448_022.nt.1	791	791.0	Lung array	457-516
DEX0448_022.nt.1	855	855.0	Lung array	285-344
DEX0448_022.nt.1	804	804.0	Lung array	438-497
DEX0448_022.nt.1	32151	32151.0	Breast array	163-222
DEX0448_022.nt.1	803	803.0	Lung array	441-500
DEX0448_022.nt.1	29143	29143.01	Prostate2 array	259-318
DEX0448_022.nt.1	32150	32150.0	Breast array	203-262
DEX0448_022.nt.2	804	804.0	Lung array	333-392
DEX0448_022.nt.2	37415	37415.0	Colon array	347-406
DEX0448_022.nt.2	791	791.0	Lung array	352-411
DEX0448_022.nt.2	32151	32151.0	Breast array	58-117
DEX0448_022.nt.2	36878	36878.0	Colon array	397-456
DEX0448_022.nt.2	792	792.0	Lung array	347-406
DEX0448_022.nt.2	29143	29143.01	Prostate2 array	154-213
DEX0448_022.nt.2	803	803.0	Lung array	336-395
DEX0448_022.nt.2	855	855.0	Lung array	180-239
DEX0448_023.nt.1	22297	22297.0	Colon array	1117-1176
DEX0448_024.nt.1	19607	19607.0	Colon array	199-258
DEX0448_025.nt.1	40034	40034.0	Colon array	583-642
DEX0448_025.nt.1	40033	40033.0	Colon array	784-843
DEX0448_026.nt.1	78479	78479.0	Multi-Cancer array	1695-1754
DEX0448_026.nt.1	41284	41284.0	Colon array	2169-2228
DEX0448_026.nt.1	41210	41210.0	Colon array	1892-1951
DEX0448_026.nt.2	78479	78479.0	Multi-Cancer array	1071-1130
DEX0448_026.nt.2	41284	41284.0	Colon array	1545-1604
DEX0448_027.nt.1	15836	15836.0	Breast array	5393-5452
DEX0448_027.nt.1	32137	32137.0	Breast array	5280-5339

DEX0448 027.nt.1	15311	15311.0	Breast array	4539-4598
DEX0448 027.nt.1	15846	15846.0	Breast array	5624-5683
DEX0448 027.nt.1	15845	15845.0	Breast array	5871-5930
DEX0448 027.nt.1	15310	15310.0	Breast array	4588-4647
DEX0448 027.nt.1	32136	32136.0	Breast array	5363-5422
DEX0448 027.nt.1	15834	15834.0	Breast array	3574-3633
DEX0448 027.nt.1	9623	9623.0	Colon array	5423-5482
DEX0448 027.nt.1	15835	15835.0	Breast array	5423-5482
DEX0448 027.nt.1	9622	9622.0	Colon array	5527-5586
DEX0448 027.nt.1	15833	15833.0	Breast array	3594-3653
DEX0448 027.nt.2	15845	15845.0	Breast array	5739-5798
DEX0448 027.nt.2	15833	15833.0	Breast array	3462-3521
DEX0448 027.nt.2	9622	9622.0	Colon array	5395-5454
DEX0448 027.nt.2	15311	15311.0	Breast array	4407-4466
DEX0448 027.nt.2	15846	15846.0	Breast array	5492-5551
DEX0448 027.nt.2	15836	15836.0	Breast array	5261-5320
DEX0448 027.nt.2	15310	15310.0	Breast array	4456-4515
DEX0448 027.nt.2	15835	15835.0	Breast array	5291-5350
DEX0448 027.nt.2	32136	32136.0	Breast array	5231-5290
DEX0448 027.nt.2	9623	9623.0	Colon array	5291-5350
DEX0448 027.nt.2	15834	15834.0	Breast array	3442-3501
DEX0448 027.nt.2	32137	32137.0	Breast array	5148-5207
DEX0448 027.nt.3	15836	15836.0	Breast array	7132-7191
DEX0448 027.nt.3	15833	15833.0	Breast array	4013-4072
DEX0448 027.nt.3	9622	9622.0	Colon array	7266-7325
DEX0448 027.nt.3	32137	32137.0	Breast array	7019-7078
DEX0448 027.nt.3	9623	9623.0	Colon array	7162-7221
DEX0448 027.nt.3	15835	15835.0	Breast array	7162-7221
DEX0448 027.nt.3	32136	32136.0	Breast array	7102-7161
DEX0448 027.nt.3	15310	15310.0	Breast array	6327-6386
DEX0448 027.nt.3	15834	15834.0	Breast array	3993-4052
DEX0448 027.nt.3	15846	15846.0	Breast array	7363-7422
DEX0448 027.nt.3	15311	15311.0	Breast array	6278-6337
DEX0448 027.nt.3	15845	15845.0	Breast array	7610-7669
DEX0448 027.nt.4	15836	15836.0	Breast array	5124-5183
DEX0448 027.nt.4	15311	15311.0	Breast array	4270-4329
DEX0448 027.nt.4	32136	32136.0	Breast array	5094-5153
DEX0448 027.nt.4	15833	15833.0	Breast array	3325-3384
DEX0448 027.nt.4	15834	15834.0	Breast array	3305-3364
DEX0448 027.nt.4	9623	9623.0	Colon array	5154-5213
DEX0448 027.nt.4	15835	15835.0	Breast array	5154-5213
DEX0448 027.nt.4	9622	9622.0	Colon array	5258-5317
DEX0448 027.nt.4	32137	32137.0	Breast array	5011-5070
DEX0448 027.nt.4	15310	15310.0	Breast array	4319-4378
DEX0448 027.nt.5	15835	15835.0	Breast array	5154-5213
DEX0448 027.nt.5	9623	9623.0	Colon array	5154-5213
DEX0448 027.nt.5	15836	15836.0	Breast array	5124-5183

DEX0448_027.nt.5	15834	15834.0	Breast array	3305-3364
DEX0448_027.nt.5	9622	9622.0	Colon array	5258-5317
DEX0448_027.nt.5	15833	15833.0	Breast array	3325-3384
DEX0448_027.nt.5	32136	32136.0	Breast array	5094-5153
DEX0448_027.nt.5	15311	15311.0	Breast array	4270-4329
DEX0448_027.nt.5	32137	32137.0	Breast array	5011-5070
DEX0448_027.nt.6	15846	15846.0	Breast array	1797-1856
DEX0448_027.nt.6	15835	15835.0	Breast array	1596-1655
DEX0448_027.nt.6	9622	9622.0	Colon array	1700-1759
DEX0448_027.nt.6	15845	15845.0	Breast array	2044-2103
DEX0448_027.nt.6	9623	9623.0	Colon array	1596-1655
DEX0448_027.nt.6	15310	15310.0	Breast array	761-820
DEX0448_027.nt.6	15836	15836.0	Breast array	1566-1625
DEX0448_027.nt.6	32137	32137.0	Breast array	1453-1512
DEX0448_027.nt.6	32136	32136.0	Breast array	1536-1595
DEX0448_027.nt.6	15311	15311.0	Breast array	712-771
DEX0448_028.nt.1	5305	5305.0	Multi-Cancer array	679-738
DEX0448_028.nt.1	1263	1263.0	Lung array	653-712
DEX0448_028.nt.1	1265	1265.0	Lung array	648-707
DEX0448_028.nt.1	1262	1262.0	Lung array	674-733
DEX0448_028.nt.1	5306	5306.0	Multi-Cancer array	648-707
DEX0448_028.nt.1	1320	1320.0	Lung array	988-1047
DEX0448_028.nt.1	1264	1264.0	Lung array	679-738
DEX0448_028.nt.2	1262	1262.0	Lung array	423-482
DEX0448_028.nt.2	1264	1264.0	Lung array	428-487
DEX0448_028.nt.2	5305	5305.0	Multi-Cancer array	428-487
DEX0448_028.nt.2	1263	1263.0	Lung array	402-461
DEX0448_028.nt.2	1320	1320.0	Lung array	812-871
DEX0448_028.nt.2	1265	1265.0	Lung array	397-456
DEX0448_028.nt.2	5306	5306.0	Multi-Cancer array	397-456
DEX0448_029.nt.1	38381	38381.0	Colon array	1892-1951
DEX0448_029.nt.1	28733	28733.0	Colon array	2187-2246
DEX0448_029.nt.1	38383	38383.0	Colon array	2546-2605
DEX0448_029.nt.1	38384	38384.0	Colon array	2526-2585
DEX0448_029.nt.1	38382	38382.0	Colon array	1872-1931
DEX0448_029.nt.1	28734	28734.0	Colon array	2147-2206
DEX0448_029.nt.2	38383	38383.0	Colon array	2101-2160
DEX0448_029.nt.2	28733	28733.0	Colon array	1742-1801
DEX0448_029.nt.2	38381	38381.0	Colon array	1447-1506
DEX0448_029.nt.2	38384	38384.0	Colon array	2081-2140
DEX0448_029.nt.2	38382	38382.0	Colon array	1427-1486
DEX0448_029.nt.2	28734	28734.0	Colon array	1702-1761
DEX0448_029.nt.3	28733	28733.0	Colon array	1745-1804
DEX0448_029.nt.3	28734	28734.0	Colon array	1705-1764
DEX0448_029.nt.3	38382	38382.0	Colon array	1430-1489
DEX0448_029.nt.3	38384	38384.0	Colon array	2084-2143
DEX0448_029.nt.3	38381	38381.0	Colon array	1450-1509

DEX0448_029.nt.3	38383	38383.0	Colon array	2104-2163
DEX0448_029.nt.4	28734	28734.0	Colon array	2219-2278
DEX0448_029.nt.4	38382	38382.0	Colon array	1944-2003
DEX0448_029.nt.4	38384	38384.0	Colon array	2598-2657
DEX0448_029.nt.4	38381	38381.0	Colon array	1964-2023
DEX0448_029.nt.4	28733	28733.0	Colon array	2259-2318
DEX0448_029.nt.5	38383	38383.0	Colon array	1910-1969
DEX0448_029.nt.5	38381	38381.0	Colon array	1256-1315
DEX0448_029.nt.5	38382	38382.0	Colon array	1236-1295
DEX0448_029.nt.5	38384	38384.0	Colon array	1890-1949
DEX0448_029.nt.5	28734	28734.0	Colon array	1511-1570
DEX0448_029.nt.6	38383	38383.0	Colon array	1944-2003
DEX0448_029.nt.6	28733	28733.0	Colon array	1585-1644
DEX0448_029.nt.6	38381	38381.0	Colon array	1388-1447
DEX0448_029.nt.6	38384	38384.0	Colon array	1924-1983
DEX0448_029.nt.6	38382	38382.0	Colon array	1368-1427
DEX0448_029.nt.6	28734	28734.0	Colon array	1545-1604
DEX0448_029.nt.7	28733	28733.0	Colon array	1332-1391
DEX0448_029.nt.7	28734	28734.0	Colon array	1292-1351
DEX0448_029.nt.7	38382	38382.0	Colon array	1017-1076
DEX0448_029.nt.7	38384	38384.0	Colon array	1671-1730
DEX0448_029.nt.7	38381	38381.0	Colon array	1037-1096
DEX0448_029.nt.7	38383	38383.0	Colon array	1691-1750
DEX0448_030.nt.1	29604	29604.0	Colon array	1800-1859
DEX0448_030.nt.1	40868	40868.0	Colon array	1413-1472
DEX0448_030.nt.1	40867	40867.0	Colon array	1536-1595
DEX0448_030.nt.1	29603	29603.0	Colon array	1840-1899
DEX0448_032.nt.1	37239	37239.0	Colon array	258-317
DEX0448_032.nt.1	36179	36179.0	Colon array	531-590
DEX0448_032.nt.1	37227	37227.0	Colon array	663-722
DEX0448_032.nt.1	37240	37240.0	Colon array	238-297
DEX0448_032.nt.1	36184	36184.0	Colon array	157-216
DEX0448_032.nt.1	36180	36180.0	Colon array	490-549
DEX0448_032.nt.1	36182	36182.0	Colon array	663-722
DEX0448_032.nt.1	37228	37228.0	Colon array	633-692
DEX0448_033.nt.1	22294	22294.0	Breast array	1462-1521
DEX0448_034.nt.1	22546	22546.0	Breast array	236-295
DEX0448_034.nt.1	21619	21619.01	Ovarian array	117-176
DEX0448_034.nt.1	5353	5353.0	Lung array	267-326
DEX0448_034.nt.1	22545	22545.0	Breast array	266-325
DEX0448_034.nt.1	5354	5354.0	Multi-Cancer array	256-315
DEX0448_034.nt.1	21613	21613.02	Ovarian array	400-459
DEX0448_035.nt.1	3729	3729.0	Lung array	3628-3687
DEX0448_035.nt.1	23863	23863.0	Colon array	4553-4612
DEX0448_035.nt.1	23862	23862.0	Colon array	4617-4676
DEX0448_035.nt.1	40309	40309.0	Breast array	4890-4949
DEX0448_035.nt.2	23863	23863.0	Colon array	4372-4431

DEX0448_035.nt.2	3729	3729.0	Lung array	3447-3506
DEX0448_035.nt.2	40309	40309.0	Breast array	4709-4768
DEX0448_035.nt.2	23862	23862.0	Colon array	4436-4495
DEX0448_035.nt.3	3729	3729.0	Lung array	3358-3417
DEX0448_035.nt.3	23863	23863.0	Colon array	4283-4342
DEX0448_035.nt.3	40309	40309.0	Breast array	4620-4679
DEX0448_035.nt.3	23862	23862.0	Colon array	4347-4406
DEX0448_035.nt.4	40309	40309.0	Breast array	4106-4165
DEX0448_035.nt.4	3729	3729.0	Lung array	2844-2903
DEX0448_035.nt.4	23863	23863.0	Colon array	3769-3828
DEX0448_036.nt.1	36812	36812.0	Colon array	318-377
DEX0448_036.nt.1	29533	29533.0	Colon array	660-719
DEX0448_036.nt.1	36811	36811.0	Colon array	348-407
DEX0448_036.nt.1	29534	29534.0	Colon array	640-699
DEX0448_036.nt.1	29532	29532.0	Colon array	1444-1503
DEX0448_036.nt.1	29540	29540.0	Colon array	1265-1324
DEX0448_036.nt.2	29539	29539.0	Colon array	983-1042
DEX0448_036.nt.2	29532	29532.0	Colon array	1250-1309
DEX0448_036.nt.3	29539	29539.0	Colon array	289-348
DEX0448_036.nt.3	29532	29532.0	Colon array	556-615
DEX0448_036.nt.3	29540	29540.0	Colon array	54-113
DEX0448_037.nt.1	36341	36341.0	Colon array	645-704
DEX0448_037.nt.1	36342	36342.0	Colon array	565-624
DEX0448_037.nt.2	36341	36341.0	Colon array	738-797
DEX0448_037.nt.2	36342	36342.0	Colon array	658-717
DEX0448_037.nt.3	36341	36341.0	Colon array	1583-1642
DEX0448_037.nt.3	36342	36342.0	Colon array	1503-1562
DEX0448_037.nt.4	36341	36341.0	Colon array	774-833
DEX0448_037.nt.4	36342	36342.0	Colon array	694-753
DEX0448_037.nt.5	36341	36341.0	Colon array	774-833
DEX0448_037.nt.6	36342	36342.0	Colon array	694-753
DEX0448_038.nt.1	20896	20896.0	Colon array	40-99
DEX0448_038.nt.1	20895	20895.0	Colon array	80-139
DEX0448_039.nt.1	38855	38855.0	Colon array	913-972
DEX0448_040.nt.1	40756	40756.0	Colon array	191-250
DEX0448_040.nt.1	31347	31347.0	Colon array	237-296
DEX0448_040.nt.1	40755	40755.0	Colon array	237-296
DEX0448_040.nt.2	40756	40756.0	Colon array	487-546
DEX0448_040.nt.2	7435	7435.0	Lung array	259-318
DEX0448_040.nt.2	31347	31347.0	Colon array	533-592
DEX0448_040.nt.2	7434	7434.0	Lung array	299-358
DEX0448_040.nt.2	31346	31346.0	Colon array	665-724
DEX0448_040.nt.2	40755	40755.0	Colon array	533-592
DEX0448_040.nt.3	7435	7435.0	Lung array	268-327
DEX0448_040.nt.3	40755	40755.0	Colon array	542-601
DEX0448_040.nt.3	7434	7434.0	Lung array	308-367
DEX0448_040.nt.3	31347	31347.0	Colon array	542-601

DEX0448_040.nt.3	31346	31346.0	Colon array	674-733
DEX0448_040.nt.3	40756	40756.0	Colon array	496-555
DEX0448_040.nt.4	7434	7434.0	Lung array	249-308
DEX0448_040.nt.4	31346	31346.0	Colon array	615-674
DEX0448_040.nt.4	31347	31347.0	Colon array	483-542
DEX0448_040.nt.4	7435	7435.0	Lung array	209-268
DEX0448_040.nt.4	40755	40755.0	Colon array	483-542
DEX0448_040.nt.5	7435	7435.0	Lung array	155-214
DEX0448_040.nt.5	31347	31347.0	Colon array	429-488
DEX0448_040.nt.5	40756	40756.0	Colon array	383-442
DEX0448_040.nt.5	7434	7434.0	Lung array	195-254
DEX0448_040.nt.5	31346	31346.0	Colon array	561-620
DEX0448_040.nt.5	40755	40755.0	Colon array	429-488
DEX0448_040.nt.6	35943	35943.0	Colon array	822-881
DEX0448_040.nt.6	40756	40756.0	Colon array	1684-1743
DEX0448_040.nt.6	31347	31347.0	Colon array	1730-1789
DEX0448_040.nt.6	35944	35944.0	Colon array	779-838
DEX0448_040.nt.6	7434	7434.0	Lung array	1496-1555
DEX0448_040.nt.6	7435	7435.0	Lung array	1456-1515
DEX0448_040.nt.6	31346	31346.0	Colon array	1862-1921
DEX0448_040.nt.6	40755	40755.0	Colon array	1730-1789
DEX0448_040.nt.7	31346	31346.0	Colon array	310-369
DEX0448_040.nt.7	40756	40756.0	Colon array	132-191
DEX0448_040.nt.7	31347	31347.0	Colon array	178-237
DEX0448_041.nt.1	32724	32724.03	Prostate2 array	1976-2035
DEX0448_041.nt.1	32720	32720.01	Prostate2 array	2016-2075
DEX0448_041.nt.1	28451	28451.01	Prostate1 array	2240-2299
DEX0448_041.nt.1	32216	32216.0	Colon array	1632-1691
DEX0448_041.nt.1	32718	32718.01	Prostate2 array	2016-2075
DEX0448_041.nt.1	32716	32716.03	Prostate2 array	168-227
DEX0448_041.nt.2	32216	32216.0	Colon array	1632-1691
DEX0448_041.nt.2	32716	32716.03	Prostate2 array	168-227
DEX0448_042.nt.1	7412	7412.0	Lung array	1245-1304
DEX0448_042.nt.1	29231	29231.0	Colon array	1278-1337
DEX0448_042.nt.1	29272	29272.0	Colon array	1587-1646
DEX0448_042.nt.1	7413	7413.0	Lung array	1158-1217
DEX0448_042.nt.1	29232	29232.0	Colon array	1229-1288
DEX0448_042.nt.1	29290	29290.0	Colon array	1349-1408
DEX0448_042.nt.1	29271	29271.0	Colon array	1627-1686
DEX0448_043.nt.1	9043	9043.0	Colon array	181-240
DEX0448_043.nt.1	22586	22586.0	Breast array	117-176
DEX0448_043.nt.2	966	966.0	Lung array	716-775
DEX0448_043.nt.2	9043	9043.0	Colon array	240-299
DEX0448_043.nt.2	943	943.0	Lung array	942-1001
DEX0448_043.nt.2	960	960.0	Lung array	436-495
DEX0448_043.nt.2	22585	22585.0	Breast array	222-281
DEX0448_043.nt.2	939	939.0	Lung array	658-717

DEX0448_043.nt.2	944	944.0	Lung array	922-981
DEX0448_043.nt.2	22586	22586.0	Breast array	178-237
DEX0448_043.nt.2	940	940.0	Lung array	521-580
DEX0448_043.nt.2	946	946.0	Lung array	1076-1135
DEX0448_044.nt.1	36803	36803.0	Colon array	433-492
DEX0448_044.nt.1	42013	42013.0	Multi-Cancer array	628-687
DEX0448_044.nt.2	36803	36803.0	Colon array	827-886

Example 2b: Relative Quantitation of Gene Expression

Real-Time quantitative PCR with fluorescent Taqman[®] probes is a quantitation
 5 detection system utilizing the 5'-3' nuclease activity of Taq DNA polymerase. The
 method uses an internal fluorescent oligonucleotide probe (Taqman[®]) labeled with a 5'
 reporter dye and a downstream, 3' quencher dye. During PCR, the 5'-3' nuclease activity
 of Taq DNA polymerase releases the reporter, whose fluorescence can then be detected by
 the laser detector of the Model 7700 Sequence Detection System (PE Applied Biosystems,
 10 Foster City, CA, USA). Amplification of an endogenous control is used to standardize the
 amount of sample RNA added to the reaction and normalize for Reverse Transcriptase
 (RT) efficiency. Either cyclophilin, glyceraldehyde-3-phosphate dehydrogenase
 (GAPDH), ATPase, or 18S ribosomal RNA (rRNA) is used as this endogenous control.
 To calculate relative quantitation between all the samples studied, the target RNA levels
 15 for one sample were used as the basis for comparative results (calibrator). Quantitation
 relative to the "calibrator" can be obtained using the comparative method (User Bulletin
 #2: ABI PRISM 7700 Sequence Detection System).

The tissue distribution and the level of the target gene are evaluated for every
 sample in normal and cancer tissues. Total RNA is extracted from normal tissues, cancer
 20 tissues, and from cancers and the corresponding matched adjacent tissues. Subsequently,
 first strand cDNA is prepared with reverse transcriptase and the polymerase chain reaction
 is done using primers and Taqman[®] probes specific to each target gene. The results are
 analyzed using the ABI PRISM 7700 Sequence Detector. The absolute numbers are
 relative levels of expression of the target gene in a particular tissue compared to the
 25 calibrator tissue.

One of ordinary skill can design appropriate primers. The relative levels of
 expression of the CSNA versus normal tissues and other cancer tissues can then be
 determined. All the values are compared to the calibrator. Normal RNA samples are

commercially available pools, originated by pooling samples of a particular tissue from different individuals.

The relative levels of expression of the CSNA in pairs of matched samples may also be determined. A matched pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual. All the values are compared to the calibrator.

In the analysis of matching samples, the CSNAs show a high degree of tissue specificity for the tissue of interest. These results confirm the tissue specificity results obtained with normal pooled samples. Further, the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual are compared. This comparison provides an indication of specificity for the cancer state (*e.g.* higher levels of mRNA expression in the cancer sample compared to the normal adjacent).

Information on the samples tested in the QPCR experiments below include the Sample ID (Smpl ID), Organ, Tissue Type (Tiss Type), Diagnosis (DIAG), Disease Detail, and Stage or Grade (STG or GRD) in following table.

Sample ID	ORGAN	TISS TYPE	DIAG	DISEASE DETAIL	STG or GRD
AS12	Colon	CAN		T	StageB
AS12	Colon	NAT		NL	
AS46	Colon	CAN		malignant	T3N1MX
AS46	Colon	NAT		NAT	
B34	Colon	CAN	Adenocarcinoma		
B34	Colon	NAT		NAT	
C9XR	Colon	CAN		Rectum Cancer	Stage D
C9XR	Colon	NAT		NAT	
CM67	Colon	CAN	Adenocarcinoma	Adenocarcinoma of cecum, Moderately differentiated	Stage II
CM67	Colon	NAT		NAT	
TX89	Colon	CAN	Adenocarcinoma	Adenocarcinoma of Transverse Colon	Stage IV
TX89	Colon	NAT		NAT	
AS43	Colon	CAN	Adenocarcinoma	malignant	
AS43	Colon	NAT	Adenocarcinoma	NAT	
AS98	Colon	CAN	Adenocarcinoma	Moderately to poorly differentiated adenocarcinoma	Duke's C
AS98	Colon	NAT		NAT	

RS53	Colon	CAN	Adenocarcinoma	moderately differentiated adenocarcinoma	
RS53	Colon	NAT	Adenocarcinoma	NAT	
RC01	Colon	CAN	Cancer		Stage IV
RC01	Colon	NAT		NAT	
SG27	Colon	CAN		malig	Stage B
SG27	Colon	NAT		NAT	
DC19	Colon	CAN		T	Stage B
DC19	Colon	NAT		NL	
401C	Colon	CAN	Adenocarcinoma	Adenocarcinoma of ascending colon and cecum	Stage III
401C	Colon	NAT		NAT	
CM12	Colon	CAN		T	Stage D
CM12	Colon	NAT	Adenocarcinoma	Nat	
TX01	Colon	CAN	Adenocarcinoma	Moderately differentiated adenocarcinoma of cecum	Stage II; T3NoMo
TX01	Colon	NAT		NAT	
030B	Urinary Bladder	CAN	Carcinoma	invasive Carcinoma, poorly differentiated	Stage III, Grade 3
030B	Urinary Bladder	NAT		NAT	
TR17	Urinary Bladder	CAN	Carcinoma	transitional cell carcinoma	Stage II/Grade III
TR17	Urinary Bladder	NAT		NAT	
520B	Urinary Bladder	CAN	Sarcomatoid transitional cell carcinoma	Sarcomatoid transitional cell carcinoma	
520B	Urinary Bladder	NAT		NAT	
KS52	Cervix	CAN	Squamous cell carcinoma	Keratinizing Squamous Cell Carcinoma	IIIB, well diff. G1; T3bNxM0
KS52	Cervix	NAT		NAT	
NK23	Cervix	CAN		Nonkeratinizing Large Cell	FIGO IIIB, undiff. G4; T3bNxM0
NK23	Cervix	NAT		NAT	
NKS54	Cervix	CAN	Squamous cell carcinoma	Nonkeratinizing Squamous Cell Carcinoma	IIB, mod diff. G2; T2bNxM0
NKS54	Cervix	NAT		NAT	

NKS55	Cervix	CAN	Squamous cell carcinoma	Nonkeratinizing Squamous Cell Carcinoma	IIIB, Mod diff. G2; T3bNxM0
NKS55	Cervix	NAT		NAT	
NKS81	Cervix	CAN	Squamous cell carcinoma	large cell nonkeratinizing sq carc, IIB, moderately diff	IIB
NKS81	Cervix	NAT		NAT	
10479	Endometrium	CAN		malignant mixed mullerian tumor	T?, Nx, M1
10479	Endometrium	NAT		NAT	
28XA	Endometrium	CAN	Endometrial adenocarcinoma	malignant	II/III
28XA	Endometrium	NAT		NAT	II/III
8XA	Endometrium	CAN	mod. diff, invasive, squamous differentiation, FIGO-II		
8XA	Endometrium	NAT		NAT	
106XD	Kidney	CAN	Renal cell carcinoma	renal cell carcinoma, clear cell, localized	3
106XD	Kidney	NAT		NL	
107XD	Kidney	CAN	Renal cell carcinoma	renal cell carcinoma, clear cell, with metastatic	G III
107XD	Kidney	NAT		NL	
109XD	Kidney	CAN		Malignant	G III
109XD	Kidney	NAT		NL	
10XD	Kidney	CAN	Renal cell carcinoma	renal cell carcinoma, clear cell, localized, grade 2-3	3
10XD	Kidney	NAT		NL	
22K	Kidney	CAN	Renal cell carcinoma	Renal cell carcinoma	G2, Mod. Diff.
22K	Kidney	NAT		NAT	
15XA	Liver	CAN		Sarcoma, Retroperitoneal Tumor	Grade-2
15XA	Liver	NAT		CA	St. I, G4
174L	Liver	CAN	Hepatocellular carcinoma	Moderate to well differentiated hepatocellular carcinoma	
174L	Liver	NAT	Hepatocellular carcinoma	NAT	

187L	Liver	CAN	Adenocarcinoma	Metastatic Adenocarcinoma	Liver (Gallbladder)
187L	Liver	NAT		NAT	
205L	Lung	CAN	Adenocarcinoma	poorly differentiated adenocarcinoma	T2, N1, Mx
205L	Lung	NAT		NAT	
315L	Lung	CAN	Squamous cell carcinoma		
315L	Lung	NAT	Adenocarcinoma	NAT	
507L	Lung	CAN	Bronchioloalveolar carcinoma	bronchioalveolar carcinoma	Stage IB, G1, well diff.
507L	Lung	NAT		NAT	
528L	Lung	CAN	Adenocarcinoma	Adenocarcinoma	St. IV, T2N0M1, infiltrating poorly diff.
528L	Lung	NAT		NAT	
8837L	Lung	CAN	Squamous cell carcinoma	Squamous cell carcinoma	T2, N0, M0
8837L	Lung	NAT		NAT	
AC11	Lung	CAN	Adenocarcinoma	poorly differentiated adenocarcinoma	T2, N2, M1
AC11	Lung	NAT		NAT	
AC39	Lung	CAN	Adenocarcinoma	intermediate grade adnocarcinoma	T2, N2, Mx
AC39	Lung	NAT		NAT	
SQ80	Lung	CAN	Squamous cell carcinoma	poorly differentiated squamous cell carcinoma	T1, N1, M0
SQ80	Lung	NAT		NAT	
SQ81	Lung	CAN	Squamous cell carcinoma	poorly differentiated squamous carcinoma	T3, N1, Mx
SQ81	Lung	NAT		NAT	
19DN	Mammary	CAN	Invasive ductal carcinoma	Invasive ductal carcinoma	G3, Stage IIA; T2N0M0
19DN	Mammary	NAT		NAT	
42DN	Mammary	CAN	Invasive ductal carcinoma	Invasive Ductal Carcinoma	T3aN1M0 IIIA, G3
42DN	Mammary	NAT		NAT	
517	Mammary	CAN	Infiltrating ductal carcinoma	Infiltrating ductal carcinoma	St. IIA, G3
517	Mammary	NAT		NAT	

781M	Mammary	CAN	Invasive ductal carcinoma		Architectural grade-3/3, Nuclear grade-3/3
781M	Mammary	NAT		NAT	
869M	Mammary	CAN	Invasive carcinoma	Invasive Carcinoma	Stage IIA G1;T2NoMo
869M	Mammary	NAT		NAT	
976M	Mammary	CAN	Invasive ductal carcinoma	Invasive Ductal Carcinoma	T2N1M0 (Stage 2B Grade 2-3)
976M	Mammary	NAT		NAT	
S570	Mammary	CAN	Carcinoma	Carcinoma	Stage IIA;T1N1Mo
S570	Mammary	NAT		NAT	
S699	Mammary	CAN	Invasive lobular carcinoma	Invasive Lobular Carcinoma	Stage IIB G1;T2N1Mo
S699	Mammary	NAT		NAT	
S997	Mammary	CAN	Invasive ductal carcinoma	Invasive Ductal Carcinoma	Stage IIB G3; T2N1Mo
S997	Mammary	NAT		NAT	
G021	Ovary	CAN	Carcinoma	St. IIIC, poorly diff.	Stage- IIIC, poorly diff.
G021	Ovary	NAT		NAT	
10050	Ovary	CAN		papillary serous and endometrioid ovarian carcinoma, concurrent metastatic breast cancer	3
10400	Ovary	CAN		papillary serous adeno, metastatic	
1050	Ovary	CAN		Papillary Serous Carcinoma with Focal Mucinous Differentiation	Stage IC G0; T1cN0M0
130X	Ovary	CAN		Ovarian cancer	
7180	Ovary	CAN	Adenocarcinoma	malignant tumor	IIIC
A1B	Ovary	CAN	Adenocarcinoma	CA	
1230	Ovary	NRM		Normal	
18GA	Ovary	NRM		NL	
206I	Ovary	NRM		NL	
3370	Ovary	NRM		Normal	
40G	Ovary	NRM		NL	
5150	Ovary	NRM		Normal	
C004	Ovary	NRM		NL	
C177	Ovary	NRM		several fluid filled cysts	

71XL	Pancreas	CAN		villous adenoma with paneth cell metaplasia	localized
71XL	Pancreas	NAT		NL	
82XP	Pancreas	CAN		serious cystadenoma	
82XP	Pancreas	NAT		NL	
92X	Pancreas	CAN	Ductal adenocarcinoma	ductal adenocarcinoma	mod to focally poorly diff.
92X	Pancreas	NAT		NL	
23B	Prostate	CAN		Prostate tumor	Gleason's 3+4
23B	Prostate	NAT		NAT	
65XB	Prostate	CAN	Adenocarcinoma	adenocarcinoma	3+4=7
65XB	Prostate	NAT		NL	
675P	Prostate	CAN	Adenocarcinoma	adenocarcinoma	
675P	Prostate	NAT		Normal	
84XB	Prostate	CAN	Adenocarcinoma	adenocarcinoma	2+3
84XB	Prostate	NAT		NL	
958P	Prostate	CAN	Adenocarcinoma	Adenocarcinoma	T2C, NO, MX
958P	Prostate	NAT	NAT	Normal	
263C	Prostate	BPH		BPH	
276P	Prostate	BPH		BPH	
767B	Prostate	BPH		prostate BPH	
855P	Prostate	BPH		BPH	
10R	Prostate	PROST		active chronic prostatitis	T0, NO, M0
20R	Prostate	PROST		PROSTATITIS	
287S	Skin	CAN	Squamous cell carcinoma	Invasive Keratinizing Squamous Cell Carcinoma	Moderately Differentiated
287S	Skin	NAT		NAT	
39A	Skin	CAN		CA	St. II
39A	Skin	NAT		CA	St. II
669S	Skin	CAN	Melanoma	Nodular malignant melanoma	
669S	Skin	NAT		NAT	
171S	Small Intestine	CAN	Adenocarcinoma	Moderately differentiated Adenocarcinoma, invasive	
171S	Small Intestine	NAT		NAT	
20SM	Small Intestine	CAN	Adenocarcinoma	Adenocarcinoma, metastatic to lung & liver	St. IV, poorly diff.

20SM	Small Intestine	NAT		NAT	
H89	Small Intestine	CAN	Adenocarcinoma	Adenocarcinoma	80% tumor, 50% necrosis, moderately differentiated, G2-3; T3N1MX
H89	Small Intestine	NAT	Adenocarcinoma	NAT	
261S	Stomach	CAN	Signet-ring cell carcinoma	Signet-ring cell carcinoma	Stage IIIA, T3N1M0
261S	Stomach	NAT		NAT	
288S	Stomach	CAN	Adenocarcinoma	Infiltrating Adenocarcinoma	Moderately Differentiated
288S	Stomach	NAT		NAT	
AC93 or 509L	Stomach	CAN	Adenocarcinoma	Adenocarcinoma	St. IV, G4, T4N3M0, poorly diff.
AC93 or 509L	Stomach	NAT		NAT	
88S	Stomach	CAN	Adenocarcinoma	Mucinous adenocarcinoma	T3N1M0, St. IIIA
88S	Stomach	NAT		NAT	
143N	Thyroid Gland	CAN	Follicular carcinoma	Follicular Carcinoma	
143N	Thyroid Gland	NAT		NAT	
270T	Thyroid Gland	CAN		CA	
270T	Thyroid Gland	NAT		NAT	
56T	Thyroid Gland	CAN	Papillary carcinoma	Papillary Carcinoma	St. III; T4N1M0
56T	Thyroid Gland	NAT		NAT	
39X	Testes	CAN		CA	
39X	Testes	NAT		NAT	
647T	Testes	CAN	Teratocarcinoma	Teratocarcinoma	Stage IA
647T	Testes	NAT	Teratocarcinoma	NAT	
663T	Testes	CAN	Teratocarcinoma	Teratocarcinoma	
663T	Testes	NAT		NAT	
135XO	Uterus	CAN		Uterus normal	
135XO	Uterus	NAT		Uterus tumor	
85XU	Uterus	CAN		endometrial carcinoma	I
85XU	Uterus	NAT		NL	

B1	Blood	NRM		Normal	
B3	Blood	NRM		Normal	
B5	Blood	NRM		Normal	
B6	Blood	NRM		Normal	
B11	Blood	NRM		Normal	
982B	Blood	NRM		Normal	
48AD	Adrenal Gland	NRM		Normal	
10BR	Brain	NRM		Normal	
01CL	Colon	NRM		Normal	
06CV	Cervix	NRM		Normal	
01ES	Esophagus	NRM		Normal	
46HR	Heart	NRM		Normal	
00HR	Human Reference	CAN	CAN	Cancer pool	
55KD	Kidney	NRM		Normal	
89LV	Liver	NRM		Normal	
90LN	Lung	NRM		Normal	
01MA	Mammary	NRM		Normal	
84MU	Skeletal Muscle	NRM		Normal	
3APV	Ovary	NRM		Normal	
04PA	Pancreas	NRM		Normal	
59PL	Placenta	NRM		Normal	
09PR	Prostate	NRM		Normal	
21RC	Rectum	NRM		Normal	
59SM	Small Intestine	NRM		Normal	
7GSP	Spleen	NRM		Normal	
09ST	Stomach	NRM		Normal	
4GTS	Testes	NRM		Normal	
99TM	Thymus Gland	NRM		Normal	
16TR	Trachea	NRM		Normal	
57UT	Uterus	NRM		Normal	

DEX-0448_026.nt.1 (Cln259)

The relative expression level of Cln259 in various tissue samples is included below. Tissue samples include 78 pairs of matching samples, 7 non matched cancer samples, and 35 normal samples, all from various tissues annotated in the table. A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual. Of the normal samples 5 were blood samples which measured the expression levels in blood cells.

Additionally, 2 prostatitis, and 4 Benign Prostatic Hyperplasia (BPH) samples are included. All the values are compared to normal colon sample CLN01CL (calibrator).

- The table below contains the relative expression level values for the sample as compared to the calibrator. The table includes the Sample ID and expression level values for the following samples: Cancer (CAN), Normal Adjacent Tissue (NAT), Normal Tissue (NRM), Benign Prostatic Hyperplasia (BPH), and Prostatitis (PROST).

Sample ID	CAN	NAT	NRM	BPH	PROST
CLNAS12	0.74	1.22			
CLNAS46	1.13	1.92			
CLNB34	0.29	0.38			
CLNC9XR	0.52	0.94			
CLNCM67	0.98	0.97			
CLNTX89	0.93	0.37			
CLNAS43	7.88	1.10			
CLNAS98	3.38	1.22			
CLNRS53	0.33	0.82			
CLNRC01	1.25	1.47			
CLNSG27	1.12	2.04			
CLNDC19	2.80	1.19			
CLN401C	1.18	1.64			
CLNCM12	0.50	1.14			
CLNTX01	1.31	3.49			
BLD030B	0.46	0.70			
BLD520B	2.07	0.26			
BLDTR17	1.16	1.25			
CVXKS52	4.74	1.34			
CVXNK23	4.04	16.66			
CVXNKS54	2.96	0.81			
CVXNKS55	2.63	5.01			
CVXNKS81	6.61	12.74			
ENDO10479	6.39	1.17			
ENDO28XA	2.85	6.33			
ENDO8XA	1.44	0.47			
KID106XD	0.87	0.46			
KID107XD	1.42	0.91			
KID109XD	1.15	0.85			
KID10XD	0.45	0.12			
KID22K	0.80	0.28			
LNG205L	2.21	1.94			
LNG315L	1.87	1.88			
LNG507L	2.04	2.43			
LNG528L	4.22	1.49			
LNG8837L	1.89	2.24			
LNGAC11	2.36	2.71			

LNGAC39	4.81	1.51			
LNGSQ80	1.23	1.59			
LNGSQ81	4.06	2.94			
LVR15XA	1.10	1.05			
LVR174L	1.38	2.97			
LVR187L	3.46	2.44			
MAM19DN	2.90	2.01			
MAM42DN	4.40	2.31			
MAM517	13.65	2.05			
MAM781M	1.62	0.39			
MAM869M	30.36	0.86			
MAM976M	3.23	1.06			
MAMS570	1.53	3.30			
MAMS699	0.50	1.44			
MAMS997	3.39	1.32			
OVRG021	1.48	0.47			
OVR1005O	5.96				
OVR1040O	11.10				
OVR105O	1.58				
OVR130X	17.74				
OVR718O	3.55				
OVRA1B	4.85				
OVR123O			0.52		
OVR18GA			0.30		
OVR206I			0.26		
OVR337O			0.00		
OVR40G			0.39		
OVR515O			0.32		
OVRC004			3.88		
OVRC177			0.48		
PAN71XL	1.22	0.46			
PAN82XP	1.00	1.32			
PAN92X	9.27	0.85			
PRO23B	1.62	2.06			
PRO65XB	0.65	1.30			
PRO675P	2.12	1.79			
PRO84XB	1.38	1.64			
PRO958P	1.80	1.03			
PRO263C				1.38	
PRO276P				1.12	
PRO767B				1.88	
PRO855P				1.40	
PRO10R					0.87
PRO20R					0.75
SKN287S	1.13	1.47			
SKN39A	0.22	0.62			
SKN669S	0.41	1.28			

SMINT171S	2.91	1.41			
SMINT20SM	4.90	2.06			
SMINTH89	0.99	1.26			
STO261S	4.23	1.20			
STO288S	1.07	1.22			
STO88S	2.78	1.17			
THRD143N	0.93	1.37			
THRD270T	2.11	2.42			
THRD56T	4.34	1.52			
TST39X	2.50	2.82			
TST647T	4.81	2.35			
TST663T	4.19	3.04			
UTR135XO	1.14	1.18			
UTR85XU	3.40	1.57			
BLOB1			0.00		
BLOB3			0.63		
BLOB6			0.55		
BLOB11			0.58		
BLO982B			0.00		
ADR48AD			0.90		
HUMREF00H R	1.14				
BRN10BR			0.10		
CLN01CL			1.00		
ESO01ES			0.73		
HRT46HR			0.02		
KID55KD			0.12		
LVR89LV			3.12		
LNG90LN			1.69		
MAM01MA			0.53		
MSL84MU			0.02		
OVR3APV			0.48		
PAN04PA			0.75		
PLA59PL			1.66		
PRO09PR			0.87		
REC21RC			4.63		
SMINT59SM			0.30		
SPL7GSP			0.90		
STO09ST			1.59		
THYM99TM			0.20		
TRA16TR			1.12		
TST4GTS			2.95		
UTR57UT			0.77		

0.00= Negative or Not Detected

The sensitivity for Cln259 expression was calculated for the cancer samples versus normal samples. The sensitivity value indicates the percentage of cancer samples that

show levels of Cln259 at least 2 fold higher than the normal tissue or the corresponding normal adjacent form the same patient.

This specificity is an indication of the level of colon tissue specific expression of the transcript compared to all the other tissue types tested in our assay. Thus, these experiments indicate Cln259 being useful as an colon cancer diagnostic marker and/or therapeutic target.

Sensitivity and specificity data is reported in the table below.

	CLN	LNG	MAM	OVR	PRO
Sensitivity, Up vs. NAT	27%	22%	56%	0%	0%
Sensitivity, Down vs. NAT	20%	0%	22%	0%	20%
Sensitivity, Up vs. NRM	20%	33%	89%	100%	40%
Sensitivity, Down vs. NRM	20%	0%	0%	0%	0%
Specificity	3.47%	5.41%	9.19%	12.3%	4.28%

Altogether, the tissue specificity, plus the mRNA differential expression in the samples tested are believed to make Cln259 a good marker for diagnosing, monitoring, staging, imaging and treating colon cancer.

Additionally, the the tissue specificity, plus the mRNA differential expression in the samples tested are believed to make Cln259 a good marker for diagnosing, monitoring, staging, imaging and treating ovarian cancer and/or breast cancer.

Primers used for QPCR Expression Analysis of Cln259 are as follows:

(Cln259_forward): TACGCAGAGCTCATCGTCCTT (SEQ ID NO:238)

(Cln259_reverse): ACAACCACGAAGAGCCAGTCTT (SEQ ID NO:239)

(Cln259_probe): TGGCTGAGCTCTTACCTGGTTTTCAGGC (SEQ ID NO:240)

Conclusions

Altogether, the high level of tissue specificity, plus the mRNA overexpression in matched samples tested are indicative of SEQ ID NO: 1-95 being a diagnostic marker and/or a therapeutic target for cancer.

Example 3: Protein Expression

The CSNA is amplified by polymerase chain reaction (PCR) and the amplified DNA fragment encoding the CSNA is subcloned in pET-21d for expression in E. coli. In addition to the CSNA coding sequence, codons for two amino acids, Met-Ala, flanking the NH₂-terminus of the coding sequence of CSNA, and six histidines, flanking the

COOH-terminus of the coding sequence of CSNA, are incorporated to serve as initiating Met/restriction site and purification tag, respectively.

An over-expressed protein band of the appropriate molecular weight may be observed on a Coomassie blue stained polyacrylamide gel. This protein band is confirmed
5 by Western blot analysis using monoclonal antibody against 6X Histidine tag.

Large-scale purification of CSP is achieved using cell paste generated from 6-liter bacterial cultures, and purified using immobilized metal affinity chromatography (IMAC). Soluble fractions that are separated from total cell lysate were incubated with a nickel chelating resin. The column is packed and washed with five column volumes of wash
10 buffer. CSP is eluted stepwise with various concentration imidazole buffers.

Example 4: Fusion Proteins

The human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression
15 vector, preferably a mammalian expression vector. For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 2, is ligated into this
20 BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced. If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. *See, e.g., WO 96/34891.*

25 Example 5: Production of an Antibody from a Polypeptide

In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine
30 serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100, µg/ml of streptomycin. The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any

suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP20), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands *et al.*,

5 *Gastroenterology* 80: 225-232 (1981).

The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide. Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that
10 antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific
15 antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

Example 6: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

20 RNA is isolated from individual patients or from a family of individuals that have a phenotype of interest. cDNA is then generated from these RNA samples using protocols known in the art. *See*, Sambrook (2001), *supra*. The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO: 1-95. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C;
25 and 60-120 seconds at 70°C, using buffer solutions described in Sidransky *et al.*, *Science* 252(5006): 706-9 (1991). *See also* Sidransky *et al.*, *Science* 278(5340): 1054-9 (1997).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons are also determined and genomic PCR products
30 analyzed to confirm the results. PCR products harboring suspected mutations are then cloned and sequenced to validate the results of the direct sequencing. PCR products is cloned into T-tailed vectors as described in Holton *et al.*, *Nucleic Acids Res.*, 19: 1156

(1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements may also be determined. Genomic clones are nick-translated with digoxigenin deoxyuridine 5' triphosphate (Boehringer Mannheim), and
5 FISH is performed as described in Johnson *et al.*, *Methods Cell Biol.* 35: 73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C-and R-bands. Aligned images for precise mapping
10 are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. Johnson (1991). Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the
15 genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 7: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

Antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably
20 a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described above. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced. The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial
25 dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbound polypeptide. Next, 50 µl of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbound conjugate. 75 µl of
30 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution are added to each well and incubated 1 hour at room temperature.

The reaction is measured by a microtiter plate reader. A standard curve is prepared, using serial dilutions of a control sample, and polypeptide concentrations are plotted on the X-axis (log scale) and fluorescence or absorbance on the Y-axis (linear scale). The concentration of the polypeptide in the sample is calculated using the standard
5 curve.

Example 8: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide
10 alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1, $\mu\text{g/kg/day}$
15 to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day , and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 $\mu\text{g/kg/hour}$ to about 50 mg/kg/hour , either by 1-4 injections per day or by continuous
20 subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are
25 administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which
30 include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semipermeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. Sustained-release matrices include polylactides (U. S. Pat. No.3,773,919, EP 58,481, the contents of which are hereby incorporated by reference herein in their entirety), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., *Biopolymers* 22: 547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., *J. Biomed. Mater. Res.* 15: 167-277 (1981), and R. Langer, *Chem. Tech.* 12: 98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE Epstein et al., *Proc. Natl. Acad. Sci. USA* 82: 3688-3692 (1985); Hwang et al., *Proc. Natl. Acad. Sci. USA* 77: 4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324, the contents of which are hereby incorporated by reference herein in their entirety. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation.

For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides. Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably, the carrier is a parenteral carrier, more preferably, a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e. g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1 % (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container (s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 9: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form.

- 5 Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

- For example, a patient with decreased levels of a polypeptide receives a daily dose
10 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided above.

Example 10: Method of Treating Increased Levels of the Polypeptide

- Antisense or RNAi technology are used to inhibit production of a polypeptide of
15 the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

- For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the
20 treatment was well tolerated. The formulation of the antisense polynucleotide is provided above.

Example 11: Method of Treatment Using Gene Therapy

- One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a
25 subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of
30 the flask and fresh media (e. g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks. pMV-7 (Kirschmeier, P. T. et al., DNA, 7: 219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 3. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB 101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+aml2 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media.

If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

Example 12: Method of Treatment Using Gene Therapy-In Vivo

Another aspect of the present invention is using *in vivo* gene therapy methods to
5 treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide.

The polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by
10 the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, Tabata H. *et al. Cardiovasc. Res.* 35 (3): 470-479 (1997); Chao J *et al. Pharmacol. Res.* 35 (6): 517-522 (1997); Wolff J. A. *Neuromuscul. Disord.* 7 (5): 314-318 (1997), Schwartz B. *et al. Gene Ther.* 3 (5): 405-411 (1996); and Tsurumi Y. *et al. Circulation* 94 (12): 3281-3290 (1996); W0 90/11092, W0 98/11779; U. S. Patent No.
15 5,693,622; 5,705,151; 5,580,859, the contents of which are hereby incorporated by reference herein in their entirety.

The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, colon, liver, intestine and the like). The polynucleotide
20 constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be
25 delivered in liposome formulations (such as those taught in Felgner P. L. *et al. Ann. NY Acad. Sci.* 772: 126-139 (1995) and Abdallah B. *et al. Biol. Cell* 85 (1): 1-7 (1995)) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain
30 sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the

transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, colon, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. In vivo muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 $\mu\text{g/kg}$ body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to colons or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle in vivo is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice.

The results of the above experimentation in mice can be use to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

Example 13: Transgenic Animals

The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e. g., baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol. Any technique known in the art may be used to introduce the transgene (I. e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection

(Paterson et al., *Appl. Microbiol. Biotechnol.* 40: 691-698 (1994); Carver et al., *Biotechnology* 11: 1263-1270 (1993); Wright et al., *Biotechnology* 9: 830-834 (1991); and U. S. Pat. No. 4,873,191, the contents of which is hereby incorporated by reference herein in its entirety); retrovirus mediated gene transfer into germ lines (Van der Putten et al.,
5 *Proc. Natl. Acad. Sci., USA* 82: 6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., *Cell* 56: 313-321 (1989)); electroporation of cells or embryos (Lo, 1983, *Mol Cell. Biol.* 3: 1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e. g., Ulmer et al., *Science* 259: 1745 (1993); introducing nucleic acid constructs into embryonic pluripotent stem cells
10 and transferring the stem cells back into the blastocyst; and sperm mediated gene transfer (Lavitrano et al., *Cell* 57: 717-723 (1989). For a review of such techniques, see Gordon, "Transgenic Animals," *Intl. Rev. Cytol.* 115: 171-229 (1989).

Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated
15 oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell et al., *Nature* 380: 64-66 (1996); Wilmut et al., *Nature* 385: 810813 (1997)).

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, I. e., mosaic animals or chimeric. The transgene may be integrated as a single transgene
20 or as multiple copies such as in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., *Proc. Natl. Acad. Sci. USA* 89: 6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest,
25 and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and
30 disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Gu et al., *Science* 265: 103-106 (1994)). The regulatory sequences required for such a

cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished
5 by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and reverse transcriptase-PCR (rt-PCR).
10 Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding
15 strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to
20 both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited
25 to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

Example 14: Knock-Out Animals

30 Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (E. g., see Smithies et al., *Nature* 317: 230-234 (1985); Thomas & Capecchi, *Cell* 51: 503-512 (1987);

Thompson et al., Cell 5: 313-321 (1989)) Alternatively, RNAi technology may be used. For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention in vivo. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (e. g., see Thomas & Capecchi 1987 and Thompson 1989, supra). However, this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors that will be apparent to those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (e. g., knockouts) are administered to a patient in vivo. Such cells may be obtained from the patient (i.e., animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (e. g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered in vitro using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e.g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc.

The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, e. g., in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e. g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U. S. Patent No. 5,399,349; and
5 Mulligan & Wilson, U. S. Patent No. 5,460,959, the contents of which are hereby incorporated by reference herein in their entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the
10 cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of
15 polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

While preferred illustrative embodiments of the present invention are described, one skilled in the art will appreciate that the present invention can be practiced by other
20 than the described embodiments, which are presented for purposes of illustration only and not by way of limitation. The present invention is limited only by the claims that follow.

We claim:

1. An isolated nucleic acid molecule comprising:
 - (a) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 96-237;
 - 5 (b) a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95;
 - (c) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (a) or (b); or
 - (d) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (a) or (b).
- 10 2. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is a cDNA.
- 15 3. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is genomic DNA.
4. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is an RNA.
- 20 5. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is a mammalian nucleic acid molecule.
6. The nucleic acid molecule according to claim 5, wherein the nucleic acid molecule is a human nucleic acid molecule.
- 25 7. A method for determining the presence of a colon specific nucleic acid (CSNA) in a sample, comprising the steps of:
 - (a) contacting the sample with the nucleic acid molecule of SEQ ID NO: 1-95 under conditions in which the nucleic acid molecule will selectively hybridize to a colon specific nucleic acid; and
 - 30

- (b) detecting hybridization of the nucleic acid molecule to a CSNA in the sample, wherein the detection of the hybridization indicates the presence of a CSNA in the sample.
- 5 8. A vector comprising the nucleic acid molecule of claim 1.
9. A host cell comprising the vector according to claim 8.
10. A method for producing a polypeptide encoded by the nucleic acid molecule
10 according to claim 1, comprising the steps of:
- (a) providing a host cell comprising the nucleic acid molecule operably linked to one or more expression control sequences, and
- (b) incubating the host cell under conditions in which the polypeptide is produced.
- 15 11. A polypeptide encoded by the nucleic acid molecule according to claim 1.
12. An isolated polypeptide selected from the group consisting of:
- (a) a polypeptide comprising an amino acid sequence with at least 95%
20 sequence identity to of SEQ ID NO: 96-237 ; or
- (b) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95.
- 25 13. An antibody or fragment thereof that specifically binds to:
- (a) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 96-237 ; or
- (b) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule
30 comprising a nucleic acid sequence of SEQ ID NO: 1-95.
14. A method for determining the presence of a colon specific protein in a sample, comprising the steps of:

- (a) contacting the sample with a suitable reagent under conditions in which the reagent will selectively interact with the colon specific protein comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 96-237; and
- 5 (b) detecting the interaction of the reagent with a colon specific protein in the sample, wherein the detection of binding indicates the presence of a colon specific protein in the sample.
15. A method for diagnosing or monitoring the presence and metastases of colon
- 10 cancer in a patient, comprising the steps of:
- (a) determining an amount of:
- (i) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 96-237;
- (ii) a nucleic acid molecule comprising a nucleic acid sequence of SEQ
- 15 ID NO: 1-95;
- (iii) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (i) or (ii);
- (iv) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (i) or (ii);
- 20 (v) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 96-237 ; or
- (vi) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95
- 25 and;
- (b) comparing the amount of the determined nucleic acid molecule or the polypeptide in the sample of the patient to the amount of the colon specific marker in a normal control; wherein a difference in the amount of the nucleic acid molecule or the polypeptide in the sample compared to the amount of the nucleic
- 30 acid molecule or the polypeptide in the normal control is associated with the presence of colon cancer.

16. A kit for detecting a risk of cancer or presence of cancer in a patient, said kit comprising a means for determining the presence of:
- (a) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 96-237;
 - 5 (b) a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95;
 - (c) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (a) or (b); or
 - (d) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (a) or (b); or
 - 10 (e) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 96-237 ; or
 - (f) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95.
 - 15
17. A method of treating a patient with colon cancer, comprising the step of administering a composition consisting of:
- (a) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 96-237;
 - 20 (b) a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95;
 - (c) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (a) or (b);
 - 25 (d) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (a) or (b);
 - (e) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 96-237 ; or
 - (f) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95;
 - 30
- to a patient in need thereof, wherein said administration induces an immune response against the colon cancer cell expressing the nucleic acid molecule or polypeptide.

18. A vaccine comprising the polypeptide or the nucleic acid encoding the polypeptide of claim 12.

SEQUENCE LISTING

<110> diaDexus, Inc.
Macina, Roberto
Turner, Leah
Sun, Yongming

<120> Compositions, Splice Variants and Methods Relating to Colon
Specific Genes and Proteins

<130> DEX-0448

<150> US 60/431,132

<151> 2002-12-04

<150> US 60/431,144

<151> 2002-12-04

<160> 240

<170> PatentIn version 3.1

<210> 1

<211> 547

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (103)..(103)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (112)..(112)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (208)..(208)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (243)..(243)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (277)..(277)

<223> n=a, c, g or t

<400> 1

cagaagagta ttcacgcctg caatgtaggt ggccgactcc tgtgtcagga caggccacct 60

acattgcaga agagtattca cgctgcgca gccaggatag cantgagctc angacaccgg 120

ccagggacat tctcgcgagt tacagccctg aacgacgtgg aaactcgtga ttcgacgtgg 180
ccacatgccc gatgcgaagg ccctgcantg agcagagacg tgtggacccc tgcaggatgc 240
agncaggaag ctgtggagtt ggtccagtat gcttacnttt agagaagggtt cgtggagaac 300
ggagaagaaa cgtaagaagc gaagtgamgg atkwastcar rgacagaaga tgaagaggag 360
aaaagccaag aggaycmkgr agcgaagagg aagagaagga agactcgcca gccagatgcc 420
aaagatgggg attcatatga mcccctatgac ttcagtgaac cagaggagga aatgcctcaa 480
gtacacactc caaagacggc agactcacag gagaccaagg aatcccagaa agtggagttg 540
agtgaat 547

<210> 2
<211> 325
<212> DNA
<213> Homo sapien

<400> 2
ggcctggatt aatacccctc gaagacaagg aggactatgg gccgaataag gagtgtccac 60
tgtgcctttg tccaagattg tttgaatccc taagtcgaga tctcaaaagg actatggtgt 120
atacctagag gactcaggcc acactcttag aggtctcttc attattgatg acaaaggaat 180
cctaagacaa attactctga atgatcttcc tgtgggtaga tcagtggatg agacactacg 240
tttggttcaa gcattccagt acactgacaa acacggagaa gtctgccctg ccggctggaa 300
acctggtaag tgaaacaata atccc 325

<210> 3
<211> 1165
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (44)..(44)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (50)..(50)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (57)..(58)
<223> n=a, c, g or t

<220>

<221> misc_feature
<222> (60)..(61)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (66)..(66)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (68)..(68)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (70)..(74)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (76)..(76)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (78)..(78)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (80)..(81)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (83)..(84)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (87)..(87)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (95)..(95)
<223> n=a, c, g or t

<220>
<221> misc_feature

<222> (98)..(98)
 <223> n=a, c, g or t

<220>
 <221> misc_feature
 <222> (101)..(101)
 <223> n=a, c, g or t

<220>
 <221> misc_feature
 <222> (104)..(104)
 <223> n=a, c, g or t

<220>
 <221> misc_feature
 <222> (110)..(111)
 <223> n=a, c, g or t

<400> 3
 aaccgctccg gcaacgccaa ccgctccgct gcgcgcaggc tggnetgcan gctctcnnn 60
 ncagcncncn nnnnancnan nannccnctt ccaanatntc ncancctctcn ncctccctct 120
 gctgcctgct ggtgttgcc aatgcccgga gcaggccctc ttccatccc ctgtcggatg 180
 agctggtcaa ctatgtsaac aaacggaata ccacgtggca ggccgggcac aacttctaca 240
 acgtggacat gagctacttg aagaggctat gtggtacctt cctgggtggg cccaagccac 300
 cccagagagt tatgtttacc gaggacctga agctgcctgc aagcttcgat gcacgggaac 360
 aatggccaca gtgtccacc atcaaagaga tcagagacca gggctcctgt ggctcctgct 420
 gggccttcgg ggctgtggaa gccatctctg accggatctg satycacacc aatgcgcacg 480
 tcgagcgtgg aggtgtcggc ggaggacctg ctcacrtgct gtggcagcat gtgtggggac 540
 ggctgtaatg gtggctatcc tgctgaagct tggaaacttct ggacaagaaa aggcctggtt 600
 tctggtggcc atctatgaat cccatgtagg gtgcagaccg taytccatac cctccctgtg 660
 agcaccacgt cgaaggggtc ccggbcccca tgcacggggg agggagatac ccscargtgt 720
 agcaagakct gtgagcctgg gtacagaccg actayaaacw ggacamgcgy tayggataca 780
 attctacagg tstccaatag cgaggacrtc atggcygaga tctmaaaaaa cggcccgtgg 840
 agggagcttt ctctgtgtat cggacttcct ggtctagaag tcagggagtg taccaacacg 900
 tcaccggaga gatgattggg tgggcatgcc atccggatcc tgggctgggg agtggcgaag 960
 gacaccctac tgggtggttc atcgaaaca ctgactgggg tgacatgggt cttaaatacc 1020
 cgaggcagga cacgtgggac cgcacacatg ggggggggaa cggccgagca gcggaaaaat 1080
 aatgcgggcg tggcactgcg gccacaggaa aacgtaatca gcacacaaaa agaggaggca 1140

ggaacaaccc ggaaactgac gtgga

1165

<210> 4
<211> 754
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (29)..(29)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (135)..(135)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (176)..(176)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (235)..(235)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (256)..(256)
<223> n=a, c, g or t

<400> 4
tcacctccct ccacgaggac gattggcgnt ctgggccttc ccggggccct gccctcacc 60
cgatccggga tgaggagtgg ggtggccact ccccccggag tcccagggga tgggaccagg 120
agccccccag ggagnaggca ggcgggggct ggcgggccag gcggccccgg gcccgntccg 180
atagacgcca ctggacgacc tcaccccgcc gagcaccgca cgagtcaggg agcanctctc 240
ccacgaataa tgggtngaga agccgggcac tacatgcccc cgagagccc cagccgggac 300
gacctctatg accaagacaa ctctagggac attcccacgc taccacaggc gacccccatc 360
tacgacaaca ttcaggctcc tcgggagcgc cctcctgcct accccagggtc ccaccaccac 420
cgtacccggg accctcggga caacggctcc aggtccgggg acctccccta tgatgggcgg 480
ctactggagg aggctgtgag gaagaaaggg gtcggaggag aggaggatac cccacaagga 540
ggaggaggaa gaggcctact acccgcccgc gccgcccccg tactcggaga ccgactogca 600
ggcgtcccga gagcgcaggc tcaagaagaa cttggccctg agtcgggaaa gtttagtcgt 660

6

ctgatctgac gttttctacg tagctttygt atttttttwty ytaatttgaa ggcacactga 720
tgaagcatac cgtacccttc cgagttataa cgtc 754

<210> 5
<211> 1269
<212> DNA
<213> Homo sapien

<400> 5
cagccccgcgc ggggtactgtg tgagtctgcg tctgggtggga cggggacggg cgcggtggaa 60
gggtgatagg cgctgtgggt ttgagcttgg tgcgttgtgc gcgttgaggg gggttatgttg 120
ttcgggtgtcc tgtgtgtttg tatatgggat agtgtgtgct tggctgtgct gaggtgtgtg 180
ctcgggtgtg ttattggcgt ggtcagatgc gtcttggagt ttctgtgagg cgcctttgtg 240
tgtccccggg cgcgggggatg acgttgtgtt ggtcgtgggtg tgcttgtggg agccacacgt 300
gggaactgcg gttggaaaat attatagaag ggctaaatgt gggggcccat cctcccttga 360
cgggatttgt atgatgtcta gtgagggacg cgacgtatgt gggggcttgc gtttctgtc 420
gtgcatctga gagtgtacgg tgttatggcg tgcactatgc cgtttctcaa aaaagtttc 480
ttgttgaggc atagcatcag gagtctgagg tgcgtctttg gtcattaccc ttgggaatat 540
taagtgtgtt catcaggctg gtgttgttag ttagtaaaca gctgtgttaa taattagagt 600
actggttctt aagtgatatc agttattgtt agggatatgat gtaagacta ttgatttcac 660
ctatacagtg aggatatagg ttataagag atacttcac cagcataccc ttactctgcc 720
agagtagtga agctaattaa acacgtttgg tttctgaata aattgaacta aatccaaact 780
atttctctaaa atcacaggac attaaggacc aatagcatct gtgccagaga tgtactgtta 840
ttagctggga agaccaattc taacagcaaa taacagtctg agactcctca tacctcagtg 900
gttagaagca tgtctctctt gagctacagt agaggggaag ggattgttgt gtagtcaagt 960
caccatgctg aatgtacact gattccttta tgatgactgc ttaactcccc actgcctgtc 1020
ccagagaggc tttccaatgt agctcagtaa ttctgttac ttacagaca ggaaagtcc 1080
agaaaacttta agaacaaact ctgaaaacct atgagcaaat ggtgctgaat acwttttttt 1140
taaagccaca tttcattgtc ttagtcaaag caggmmtatt aastgattat taaaattcr 1200
wywtkwawa ttagcaactt caagtataac awcttwkaa ctggaataag tgtttatttt 1260
ctattaata 1269

<210> 6
<211> 1403
<212> DNA
<213> Homo sapien

7

```

<400> 6
cagcccgcgc gggactgtg tgagctgcg tctggtgga cggggacggg cgcggtggaa 60
ggttgatagg cgctgtgggt ttgagcttg tgcgttggtc gcgttgaggg gggtatgttg 120
ttcgggtgcc tgtgtgtttg tatatgggat agtgtgtgct tggctgtgct gaggtgtgtg 180
ctcgggtgtg ttattggcgt ggtcagatgc gtcttgaggt tttcgtgagg cgcctttgtg 240
tgtccccggg cgcggggatg acgttggtgt ggtcgtgggt tgcttggtgg agccacacgt 300
gggaactgcy gttggaaaat attatagaag ggctaaatgt gggggcccat cctccctga 360
cgggatttgt atgatgtcta gtgagggacg cgacgtatgt gggggcctgc gtttctgtc 420
gtgcatctga gagtgtacgg tgttatggcg tgcactatgc cgtttctcaa aaaagtttcc 480
ttgttgaggg atagcatcag gagtctgagg tcgctctttg gtcattacc ttgggaatat 540
taagtgtgtt catcaggctg gtgttgtag ttagtaaaca gctgtgttaa taattagagt 600
actggttctt aagtgatatc agttattgtt agggatgat gttaagacta ttgatttcac 660
ctatacagtg aggatatagg ttataagag atacttcac cagcataccc ttactctgcc 720
agagtagtga agctaattaa acacgtttgg tttctgaata aattgaacta aatccaaact 780
atttcctaaa atcacaggac attaaggacc aatagcatct gtgccagaga tgtactgtta 840
ttagctggga agaccaattc taacagcaaa taacagtctg agactcctca tacctcagtg 900
gttagaagca tgtctctctt gagctacagt agaggggaag ggattgttgt gtagtcaaga 960
gagacatgct tctaaccact gaggtatgag gcatgtctct cttagactac agtagagggg 1020
aagggttgtt tgtgtagtca agtcaccatg ctgaatgtac actgattcct ttatgatgac 1080
tgcttaactc cccactgcct gtcccagaga ggctttccaa ttagactcag taattcctgt 1140
tactttacag acaggaaagt tccagaaact ttaagaacaa actctgaaag acctattcag 1200
caccatttgc tcataggtct ttacagagtt gttcttaaag acctatgagc aaatgggtgt 1260
gaatacwttt tttttaaagc cacatttcat tgtcttagtc aaagcaggmt tattaastga 1320
ttatttaaaa ttcrrwytkw tawattagca acttcaagta taacawcttw kaaactggaa 1380
taagtgttta ttttctatta ata 1403

```

```

<210> 7
<211> 200
<212> DNA
<213> Homo sapien

```

```

<400> 7
ccaccggcag gttgtactca ccacctgaat gcagaggaaa atctctaact tccaaaggtc 60
ccaccaagca gttcagaaac ctaccaccag taaatgttcc aactacagaa gtctcaccaa 120
cttctcagaa aaccaccaca aaaaccacca caccaaagtc tcawgcwaca cggagtacac 180

```

ctgccaggga tcctctagaa 200

<210> 8
<211> 983
<212> DNA
<213> Homo sapien

<400> 8
ctctggggccc tgtgttttcc cgagcacctt ttctgactct ggtctggatt acatgtgtcg 60
gtatgtggca cttgtcccct tttgcccttg gtatatgca cccttctatt gtattgcgcc 120
ctttgtgccc gcactttccg gtccatgtcg gggatgatgg atcccccttc ccctttgctc 180
agttgcctcc tggggccagg ggcccctcac cacaaggggt ctggatatac agtttcataa 240
ggcctggccc tcccatgttt gcatgcctat gtacttctac gccaaacgtg tcagcccttc 300
ctcctgaagc tctctgccgt gcctccctgt tctgggagga cgggtggtgcg gtgtgactga 360
atttgggcct cttgtacagt taactctccc aggtggattt tgtggagggt agtaaaaggg 420
gcattgagac tataaagcag tagacaatcc ccacatacca tctgtagagt tggaactgca 480
ttctttcaaa gtwtkrrtaw gcatataytt gtagggctgt agacttacgg tcctawtakt 540
ctkwttacca tygettattc twgagcacwa aagtataat caattakyac asttsawaca 600
tcacctttwt tgacttgctc caagcccttt tacagctcww tggcattttc ctgcymagg 660
cctgtgaggt aactgggctc gcacctttta taccagagac ctgaggcaga tgaaatwtat 720
ttccatctag gactagaaaa acttgggtct cttaccgca gactgagagg cagaagtcag 780
cccgaatgcc tgctagtttc atggagggga aacgcaaac ctgcagttcc tgagtacctt 840
mtacaggcmc ggcccagcyt agggccgggg tggcaacacc acagcaagcc ggccccccct 900
cttttggcct tgtggataag ggagagtgga ccgttttcat cctgcctccg ttattgtaac 960
ttggagggtt acacgggtgc gtt 983

<210> 9
<211> 676
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (676)..(676)
<223> n=a, c, g or t

<400> 9
caaagacttc gactcccctg agaacggtgc ggacagcttc cagagctcag actccctcct 60
tcagtccctgg aacagccagt cgtccttgct ggatgtgcaa cgggttcctt ccttcgagag 120

9

cttsgragrt gactgcakcc wkastctmtk cctcaataag ccaacatgtm tttcaaggat 180
 tacatccaag agaggagtga cccrktggag caaggcaaac cagttatacc tgcagctgtg 240
 ctggccgggt tcacaggaag tggacctatt cagctgtggc agtwtctcct ggagctgyta 300
 tcasacaaat cctgccagtc attcatcasc tggactgkag ackgatggsa gtttaagctc 360
 gccgaccccg atgaggtggc ccgccagtgg ggaaagaggc ttagcccacg atgaactacg 420
 agaagctgaa ccggggctta cgctactatt acgactagaa catcatgcac aagacgtcgg 480
 agaaagcgct acgtgttctc gattaaatgt gcggacctca gaaacttgcg tggggtttca 540
 acgccccgag ggaactgtaa cgcgcataac ctgggcagta acaaccccg acaccggaag 600
 gtacttgacg ttccgcccgg ggaccacccc cttgaacccg gctccccaag gcttcgtggg 660
 acctgttggg tgcccn 676

<210> 10

<211> 2220

<212> DNA

<213> Homo sapien

<400> 10

gggggggaga agaggggagg gcatcatggg gccagtgccg tctgttcggc cctgggaagc 60
 tgcgctgggc tggcttgcc cctgtgtggc tgtgccaggg ccatccagga gtgttgcaact 120
 tgggacctgg aggatgggaa ggccgagagg cctttggact gctaaaccat ctggaagtca 180
 gccttttaca gacttcggca ggttcgggt ctccagggtg gatgggttca ggatggctga 240
 acctggagat agtctggagt ctctttgaag gacctgcttg gctcctgctg caaaggaaact 300
 gcaggcattt gtccttcctt tccctcccc accccactgc tgaaaaaggt tggcgggggg 360
 aaagcagctc tgcttttcac agtggtatacg tgtcaggaga cagcagaggt gctgggttga 420
 agattgcagg tggcaggccc agcccaggct gctgctcagt ggggtgcctgg ccttcatcat 480
 ctcgcccccac ctgcttcttg tgggtgtggc agtcccagct gccgtcttaa agcagatgtg 540
 cgcaacagtg tacaggcaca cctgccatag caaatccta ggatacctca gtcctccatg 600
 ttggtatgct tttcttttca acaatgtgga aaatgcagtt taccaagaag cttttttggg 660
 tttcctttta aggtattact tcaaaaaagc aagcgatgag ttgctcttg gagcgggtgt 720
 tgaggagatc tgggaggatg agacggtgct cccgatgtat gaaggccgga ttctgggcaa 780
 agtggagcgg atcgattgag ccctgggggc tggctttggg gaactgttgg agcccgaagc 840
 tcttgtgaac tgtcttggt gtgagcaact gcgacaaaac attttgaagg aaaattaaac 900
 caatgaagaa gacaaagtct aaggaagaat cggccagtgg gccttcggga gggcgggggg 960
 aggttgattt tcatgattca tgagctgggt actgactgag ataagaaaag cctgaactat 1020

10

ttattaaaaa catgaccact cttggctatt gaagatgctg cctgtatttg agagactgcc 1080
atacataata tatgacttcc tagggatctg aaatccataa actaagagaa actgtgtata 1140
gcttacctga acaggaatcc ttactgatat ttatagaaca gttgatttcc cccatcccca 1200
gtttatggat atgctgcttt aaacttgga gggggagaca ggaagtttta attgttctga 1260
ctaaacttag gagttgagct aggagtgcgt tcatggtttc ttcactaaca gaggaattat 1320
gctttgcaact acgtccctcc aagtgaagac agactgtttt agacagactt tttaaaatgg 1380
tgccctacca ttgacacatg cagaaattgg tgcgttttgt tttttttttt tcctatgctg 1440
ctctgttttg tcttaaagggt cttgagggtt gaccatgttg cgtcatcatc aacatttttg 1500
gggttgtgtt ggatgggatg atctgttgca gagggagagg cagggaaccc tgctccttcg 1560
ggccccaggt tgatcctgtg actgaggctc cccctcatgt agcctcccca ggcccagggc 1620
cctgaggcct gctagaatca ctgccgctgt gctttcgtgg aaatgacagt tccttgtttt 1680
ttttgtttct gtttttgttt tacattagtc attggaccac agccattcag gaactacccc 1740
ctgccccaca aagaaatgaa cagttgtagg gagaccagc agcaccttcc ctccacacac 1800
cttcattttg aagttcgggt ttttgtgtta agttaatctg tacattctgt ttgccattgt 1860
tacttgact atacatctgt atatagtga cgcaaaaga gtattaatcc actatctcta 1920
gtgcttgact ttaaatacgt acagtacctg tacctgcacg gtcacccgct ccgtgtgtcg 1980
ccctatatg agggctcaag ctttcccttg ttttttgaaa ggggtttatg tataaatata 2040
ttttatgcct ttttattaca agtcttgtag tcaatgactt ttgtcatgac attttgttct 2100
acttatactg taaattatgc attataaaga gttcatttaa ggaaaattac ttggtacaat 2160
aattattgta attaagagat gtagccttta ttaaaatttt atatttttca aaaaaaaaaa 2220

<210> 11
<211> 1407
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1175)..(1175)
<223> n=a, c, g or t

<400> 11
tcgacccaks gcgacwcgg aggagaagca ggagctgtcg ggaagatcag aagccagtca 60
tggaatgacca gcgcgacctt atctccaaca atgagcaact gcccatgctg ggccggcgcc 120
ctggggcccc ggagagcaag tgcagccgag gagccctgta cacaggcttt tccatcctgg 180
tgactctgct cctcgctggc caggccacca ccgcctactt cctgtaccag cagcagggcc 240

11

ggctggacaa actgacagtc acctcccaga acctgcagct ggagaacctg cgcatagaagc 300
 ttcccaagcc tcccaagcct gtgagcaaga tgcgcatggc caccctcgctg ctgatgcagg 360
 cgctgccccat gggagccctg ccccaggggc ccatgcagaa tgccaccaag tatggcaaca 420
 tgacagagga ccatgtgatg cacctgctcc agaattgctga cccctgaag gtgtaccgcg 480
 cactgaaggg gagcttcccg gagaacctga gacaccttaa gaacaccatg gagaccatag 540
 actggaaggt ctttgagagc tggatgcacc attggctcct gtttgaaatg agcaggcact 600
 ccttgagca aaagcccact gacgctccac cgaaagtact gaccaagtgc caggaagagg 660
 tcagccacat ccctggctgt ccaccgggt tcattcaggc ccaagtgcga cgagaacggc 720
 aactatctgc cactccagtg ctatggggag catcggctac tgctggtgtg tcttcccaa 780
 cggcacggag gtcccaaca ccagaagccg cgggcacat aactgcagtg agtcactgga 840
 actggaggac ccgtcttctg ggctgggtgt gaccaagcag gatctgggcc cagtcccat 900
 gtgagagcag cagaggcgggt cttcaacatc ctgccagccc cacacagcta cagctttctt 960
 gctcccttca gccccagcc cctcccatc tcccacctg tacctcatcc catgagacc 1020
 tggcgcctgg ctcttctgct acccttgagc aagacaaacc aagtcggaac agcagataac 1080
 aatgcagcaa ggcctgctg cccaatctcc atctgtcaac aggggcgtga ggtcccaggd 1140
 aagtggccaa aagcctagac agataccccg gttcnctgac avtacacagc agcctccaac 1200
 acaaggctcc aaagkaccta gggctcatgg acgagkatgg gaaggcacag ggagaagga 1260
 taaccctaca ccckagaccc caggctggac atgtgactg tctctcccc tccagccttt 1320
 ggccttggtt tttctagcct atttacmtgc aggtgagcc actcwyrctt ccttcccca 1380
 gccatcactc cccgaggaag agccaat 1407

<210> 12
 <211> 673
 <212> DNA
 <213> Homo sapien

<400> 12
 ggtcgactg gtgatgcggg ctgttgctcc ccggcctcct ggactcctgg tccskggtc 60
 tcccagcgt ggtttcgact tcagcttctt gcccagcca cctcaagaga aggctcacga 120
 tggtgccgc tactaccggg ctgatgatgc caatgtggtt cgtgaccgtg acctcgagg 180
 ggacaccacc ctcaagagcc tgagccagca gatcgagaac atccggagcc cagagggcag 240
 ccgcaagaac cccgcccga cctgccgtga cctcaagatg tgccactctg actggatgag 300
 tggcgagtag tggattgatc ccaaactaag gctgcaggct ggatgccatc ctattcggtg 360
 catgcatca atggagactg gtgagacgct gcgtgtaccc cactcagcgt cagtgtggcg 420

12

```

ccagaagaac tggtagatca gcaagaaccc caaggacaag aggcattgtct gggtcggcga 480
gagcatgacc gatggattcc agttcgagta tggcggccag ggctccgacc ctgccgatgt 540
ggccatccag ctgaccttcc tgcgcctgat gtccascgag gccttccaga acatcaccta 600
ccactgtaag aacagtgtgg cctacatgga ccaccagact ggcaacctta agaaggccct 660
gctcctccag ggc 673

```

```

<210> 13
<211> 382
<212> DNA
<213> Homo sapien

```

```

<400> 13
agcagcatgg cacttaacag agagttctct ttcattgtga tcactaccgt gacacttact 60
ttgtgcctat caggactttt tgcaatattg cgctctgtcg gctttccaat ctccagggat 120
atcatcgagc tagaccattc cctactatgg atttattttt ttccttttca aacacagtaa 180
ggaaacaatc tattactttt ttccttaaaa ggagaattta tagcactgta atacagctww 240
aaaatatttt tagaatgatg taaatagtta accttcagta gtctattaag gcattaatac 300
ttctctgaca tgcgcgtttg aggggtggagg ggtcctgaag gtgcttcacg gtctgtgatt 360
actgcttggg atgtgttctt tg 382

```

```

<210> 14
<211> 911
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (911)..(911)
<223> n=a, c, g or t

```

```

<400> 14
agcggaacgg agagaacagg aaagcgcgag gagccgvcgc caccaccagc gcagcagtc 60
tggagctgtg aggagattcg ggccgtcacc ctgcctcccc tgcgtcccg caccggccgc 120
ttctgtcctc ggaccattc caacaatctc gtaaaacatg gtggattact atgaagttct 180
aggcgtgcag agacatgcgc tctacccgag ggatattaca aaaggcatat cgggaaactg 240
gcaactgaagt ggcattccaga taaaaatcct gagaataaag aagaagcagc agcagaaaat 300
tcaagcaagt agcggaggca tatgaagtgc tgcggatgc taagaaacgg gacatctatg 360
acaaatatgg caamagaagg attaaatggg ggaggacgga ggtggaagtc attttgacag 420
tccatttgaa tttggcttca cattccgtaa ccagatgat gtcttcaggg aattttttag 480
gtggaaggga cccattttca tttgacttct ttgaagacct ttttgaggac ttctttggga 540

```

13

```
atcgaagggg tccccgagga agcagaagcc gagggacggg gtcgttttak tctgcgttca 600
gtggatttcc gtcttttgta agtggatggg cttctatgga tgcaggattt acttcattgg 660
ggtcactggg tcacgggggc ctactctat tctcttcac gtcatttggt ggtagtggca 720
tgggcaacta taaatcgata tcaacttcca ctaaattggt taatggcaga ccaatcacta 780
caaagagaat tggtgataac agtcaagaca gagtacaagt tgaagatgat ggccagttaa 840
agttcttaac tattgggtat gagcagctgc tgtgcttgga taacaagtga ttcaacgcac 900
gcgcttagct n 911
```

```
<210> 15
<211> 431
<212> DNA
<213> Homo sapien
```

```
<400> 15
ttaagakcgc kacgggcgct ttcctttcag cggagcgcgg cggcaagatg gcagtgcaaa 60
tatccaagaa gaggaagttt gtcgctgatg gcatcttcaa agctgaactg aatgagtttc 120
ttactcggga gctggctgaa gatggctact ctggagttga ggtgcgagtt acaccaacca 180
ggacagaaat cattatctta gccaccagaa cacagaatgt tcttggtgag aagggccggc 240
ggattcggga actgactgct gtagttcaga agaggtttgg ctttccagag ggcaagttag 300
agctttatgc tgaaaagggtg gccactagag gtctgtgtgc catttscca gcagagyyty 360
tgcsgtacma actcyaggas ggctcgtgc gccgcgtctt tcctatcgct gtcccgcacg 420
cccatggggc c 431
```

```
<210> 16
<211> 1047
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (7)..(7)
<223> n=a, c, g or t
```

```
<220>
<221> misc_feature
<222> (64)..(64)
<223> n=a, c, g or t
```

```
<220>
<221> misc_feature
<222> (66)..(66)
<223> n=a, c, g or t
```

<220>
 <221> misc_feature
 <222> (71)..(71)
 <223> n=a, c, g or t

<220>
 <221> misc_feature
 <222> (73)..(73)
 <223> n=a, c, g or t

<220>
 <221> misc_feature
 <222> (80)..(80)
 <223> n=a, c, g or t

<220>
 <221> misc_feature
 <222> (95)..(95)
 <223> n=a, c, g or t

<220>
 <221> misc_feature
 <222> (110)..(110)
 <223> n=a, c, g or t

<220>
 <221> misc_feature
 <222> (136)..(136)
 <223> n=a, c, g or t

<400> 16
 cccatgncct gcagaaacac ctggctgggc tgggcctgac tgaggccatt gacaagaaca 60
 aggnntnttt ntnacgcatn tcaggcaaga aggantttta cctggccan tttccacgcc 120
 accgcctttg agttgnacac agatggcaac ccctttgacc aggacatcta cgggcgcgag 180
 ggaggcgcag cccaagctg ttctacgccg accaccctt catcttcta gtgcgggaca 240
 cccaaagcgg ctccctgcta ttcatgggc gcctgggtccg gcctaagggt gacaagatgc 300
 gagacgagtt atagggcctc aggggtgcaca caggatggca ggaggcatcc aaaggctcct 360
 gagacacatg ggtgctattg gggttggggg ggaggtagg taccagcctt ggatactcca 420
 tgggggtgggg gtggaaaaac agaccggggg tcccggtgac ctgagcggac ctccagct 480
 agaattcact ccacttgac atgggccccca gatacatga tgctgagccc ggaaactcca 540
 catcctgtgg gacctgggcc atagtcattc tgctgcctt gaaagtccca gatcaagcct 600
 gcctcaatca gtattcatat ttatagccag gtaccttctc acctgtgaga ccaaattgag 660
 ctaggggggt cagccagccc tcttctgaca ctaaaacacc tcagctgcct cccagctct 720

atcccaacct ctcccaacta taaaactagg tgctgcagcc cctgggacca ggcaccccca 780
 gaatgacctg gccgcagtga ggcggattga gaaggagctc ccaggagggg cttctgggma 840
 gactctgggc aagaagcatc gtgtctggcg ttgtggggat gaactttttg ttttgtttct 900
 tcctttttta gttcttcaaa gatagggagg gaagggggaa catgagcctt tggtgctatc 960
 aatccaagaa cttatttgta catttttttt tcaataaaac ttttcccart gsaaaraaac 1020
 ccaaaaaaac cgagactagt tctctcc 1047

<210> 17
 <211> 833
 <212> DNA
 <213> Homo sapien

<400> 17
 cctgatgtcg ccacgatttc ccgctcggcc gtgggtgggtg aagctttag cctcgctcca 60
 tgaggatctt tcatgaggta tcggtcaggt cccggcccag cccggtcca acgcccggat 120
 ggctggggga gggcgtagcc ctctgatag ggccccctgt gggtgacccc ctctcccag 180
 tccctgacca tgccgcgttc gcacaaaaac cgtgaagaag gcggcccggg tcatcataga 240
 aaagtactac acgcgcctgg gcaacgactt ccacacgaac aagcgcgtgt gcgaggagat 300
 cgccattatc cccagcaaaa agctccgcaa caagatagca ggttacgtca cgcattctgat 360
 gaagcgaatt cagagaggcc cagtaagagg tatctccatc aagctgcagg aggaggagag 420
 agaaaggaga gacaattatg ttctgaggt ctgagccttg gatcaggaga ttattgaagt 480
 agatcctgac actaaggaaa tgctgaagct tttggacttc ggcagtctgt ccaaccttca 540
 ggtcactcag cctacagttg ggatgaattt caaaacgcct cggggacctg tttgaatttt 600
 ttctgtagtg ctgtattatt ttcaataaat tctgggacca ccagccttag aaacacaaga 660
 aagagaaact gggaggccta tattgcgggg gcgggaaaga ggggttgag aagatgggcc 720
 taaccggtgg tgtatcctgg ttgtcgctga cgcagaggtt tgctgtgtac tagatggggc 780
 agaatctggc cgggtcccta tagtggaggt ctgcatgaat tacaataaac gag 833

<210> 18
 <211> 1106
 <212> DNA
 <213> Homo sapien

<400> 18
 cagtaagttt ggcattgggtg cagagggagg ggtccagtgc accttttgag tccttacctc 60
 gtgggtagtt gttgccttgt gactgcccat tagggcaatt gaatagcaca ttgggtggcta 120
 tacgttggtg cacagtgtc aagtgcatag cgccctgccg gttgttcgca aggcaggagc 180

16

aactcctttt taggcaacgg gggctctctaa tgcccagagca ctgtgggctt ggtcacagga 240
ggtgcgcatg tcagcagcac ggagcctccc cgggcaggat gacttttgag ggggacacag 300
atgtctgggc aatgccaggg tcctgggaac agaggccccc agcaggacca ggagtgcggg 360
cagcgcgggc cgggggcttc tgggagccaa aggcgaggct gaggttgcaa actctggggc 420
caacatgggc cgcgttcgca ccaaaaccgt gaagaaggcg gcccggttca tcatagaaaa 480
gtactacacg cgcctgggca acgacttcca cacgaacaag cgcgtgtgcg aggagatcgc 540
cattatcccc agcaaaaagc tccgcaacaa gatagcaggt tacgtcacgc atctgatgaa 600
gcgaattcag agaggcccag taagaggtat ctccatcaag ctgcaggagg aggagagaga 660
aaggagagac aattatgttc ctgaggtctc agccttggtat caggagatta ttgaagtaga 720
tcttgacact aaggaaatgc tgaagctttt ggacttcggc agtctgtcca accttcaggt 780
cactcagcct acagttggga tgaatttcaa aacgcctcgg ggacctgtt gaattttttc 840
tgtagtgctg tattattttc aataaatstg ggacmacagc ataaaatata aaagacagag 900
agcaataaat gcaactgcaa ataatgcctg ctccaagcac gacacaaaaa acagaaccat 960
ccggcaaaac caaaacaaat agaaacaccg acatacacag caaacaagga catgaaccag 1020
caaccgcgca tgaagacaaa acaaacgcgc gacgcaaaaa gaacaaacca taccaccgaa 1080
ggatactcca caccaaccgg gcaaat 1106

<210> 19
<211> 744
<212> DNA
<213> Homo sapien

<400> 19
taacaggctt cgcctgtgct tcctgtttcc tcttttacca aggacccgcc aacatggggc 60
gcgttcgcac caaaaccgtg aagaaggcgg cccgggtcat catagaaaag tactacacgc 120
gcctgggcaa cgacttcac acgaacaagc gcgtgtgcga ggagatcgcc attatcccca 180
gcaaaaagct ccgcaacaag atagcaggtt acgtcacgca tctgatgaag cgaattcaga 240
gaggcccagt aagaggtatc tccatcaagc tgcaggagga ggagagagaa aggagagaca 300
attatgttcc tgaggtctca gccttggtatc aggagattat tgaagtagat cctgacacta 360
aggaaatgct gaagcttttg gacttcggca gtctgtccaa ctttcaggtc attcatcca 420
actgtaggct gagtgacctg aagggttgac agactgctca gcctacagtt gggatgaatt 480
tcaaaacgcc tcggggacct gtttgaattt tttctgtagt gctgtattat tttcaataaa 540
ttctgggacc accagcctta gaaacacaag aaagagaaac tgggaggcct atattgcggg 600
ggcgggaaag aggggttga gaagatgggc ctaaccgggtg gtgtatcctg gttgtcgtg 660

17

acgcagaggt ttgctgtgta ctagatgggg cagaatctgg ccgggtccct atagtggagg 720
tctgcatgaa ttacaataaa cgag 744

<210> 20
<211> 1559
<212> DNA
<213> Homo sapien

<400> 20
cttgcggctt gcggcatkgg cagtatcygc ckscatcctc ttccgtgagg cgcgctgaga 60
ccctggacyg gcctcctgar aggwtgccgg tgcgggcgcc cgcgagagg gacccgtcgc 120
catgggccgt gtgatccgtg gacagaggaa gggcgccggg tctgtgttcc gcgcgcacgt 180
gaagcacctt aaaggcgctg cgcgctcgcg cgccgtggat ttcgctgagc ggcacggcta 240
catcaagggc atcgtcaagg acatcatcca cgacccgggc cgcggcgcgc ccctcgccaa 300
gggtgtcttc cgggatccgt atcggtttaa gaagcggacg gagctgttca ttgcccga 360
gggcattcac acgggccagt ttgtgtattg cggcaagaag gccagctca acattggcaa 420
tgtgtccct gtgggcacca tgcctgaggg tacaatcgtg tgctgcctgg aggagaagcc 480
tggagaccgt ggcaagctgg cccgggcatc agggaaactat gccaccgtta tctcccacaa 540
ccctgagacc aagaagacc gtgtgaagct gccctccggc tccaagaagg ttatctctc 600
agccaacaga gctgtggttg gtgtggtggc tggaggtggc cgaattgaca aacccatctt 660
gaaggctggc cgggcgtacc acaaatataa ggcaaaggagg aactgctggc cacgagtacg 720
gggtgtggcc atgaatcctg tggagcatcc ttttgagggt ggcaaccacc agcacatcgg 780
caagccctcc accatccgca gagatgcccc tgctggccgc aaagtgggtc tcattgctgc 840
ccgccggact ggacgtctcc ggggaaccaa gactgtgcag gagaaagaga actagtgtg 900
agggcctcaa taaagtttgt gtttatgcc aaaaaaaaaa aaaaaaaaaa gaaaacaaaa 960
aaaaaaaaa aaaaaaaggc agaaaaaaaa aaagaaaaac caaagcaaaa ggaagaaaaa 1020
agaaaaaaag aaaaagaaga agaagaggca gagaggagaa caagagggcg caccacacac 1080
agccagcggg gccaaacacc accccacggc gacaacaacc gacagagaga gagaccaacc 1140
ctcaccagg cgagagaggga ggggcgcgcg gagaagaaaa aaccacaaaa aacagccgaa 1200
cacgcgcaac gaagcgacca cacacaaagc agaagaaaac accgaaaagc ctaagaagga 1260
agaggcacgc caccgcgcca ccggggcgca caacggcgca ccgaccacgc acgacagagc 1320
atagcaccaa gcagcagaag gcggagcaac catagacagc catcagaccc acagccgagg 1380
aacacgacag acgcaccacc cccaggcgcg ccgaagcata gacgaacaca cgagataaac 1440
accacagtga cggcggcgca gaggaagcac acgacacatc caacgaagga aagaacgaaa 1500

18

cacaaagcaa aacaggccac cgccagaaac aatcagtcac gccagcccca cacaccagc 1559

<210> 21

<211> 1745

<212> DNA

<213> Homo sapien

<400> 21

gcggccggcc gggcaggtag ctagaagaga ggcggtcaa agaagtagtg agataacgca 60
tatctcaggt cataggctat cccatcaccc tttatattggc gaaggrcaga gaaggaaatt 120
agtgatgatg mwggcagagg aagagaaagg tgcagcaaca rcacagcagg caagcataac 180
agcatgcatg aagcaaaagc ccaagatcgm magatgtggg ttcagatgca ggcaggcatg 240
acagcggtaa ggcataagca agaagcaaaa ctcaagcaag catcaaagca gcaaatacat 300
tgatccaggc aagcaactaa acaagaccaa gcctatttgg accagaaacc ctgatgacat 360
cacccaagag gaggatggag aattctacaa gagcctcact aatgactggg aagaccactt 420
ggcagtcaag cacttttctg tagaagggtca cgttggaatt caggscattg ctattttattc 480
ctcgtcgggc tccctttgac ctttgtgaga acaagaagaa aaagaacaac atcaaactct 540
atgtccgccg tgtgttcac atggacagct gtgatgagtt gataccagag tatctcaatt 600
ttatccgtgg tgtggttgac tctgaggatc tgcccctgaa catctccga gaaatgctcc 660
agcagagcaa aatcttgaaa gtccattcgc aaaaacattg ttaagaagtg ccttgagctc 720
ttctctagag ctggcagaag acaaggagaa ttacaagaca attcctatga gggcacttct 780
cataaaaatc tcaacgcttg gaatccacga agacatccac taaccggcgc cgctgtctg 840
agagctgctg cgctatcata cctcccagtc tggagatgag atgacatctc tgcagagta 900
tgtttctcgc atgaaggaga cacagaagtc catctattac atcactggtg agagcaaaga 960
gcaggtggcc aactcagctt ttgtggagcg agtgcggaag cggggcttcg aggtggtata 1020
tatgaccgag ccattgacg agtactgtgt gcagcagctc aaggaatttg atgggaagag 1080
cctggtctca gttaccaagg agggctctga gctgcctgag gatgaggagg agaagaagaa 1140
gatggaagag agcaaggcaa agtttgagaa cctctgcaag ctcatgaaag aaatcttaga 1200
taagaagggt gagaagggtga caatctocaa tagacttggt tcttcacctt gctgcattgt 1260
gaccagcacc tacggctgga cagccaatat ggagcggatc atgaaagccc aggcacttcg 1320
ggacaactcc accatgggct atatgatggc caaaaagcac ctggagatca accctgacca 1380
ccscatttgt ggagacgctg cggcagaagg ctgaggccga cgagaatgat aaggcagtta 1440
aggacctggt ggtgctgctg tttgaaaccg ccctgggtatc ttctggcttt tcccttgagg 1500
agtccccaga ccactccaa ccgcgatcta tcgcatgatc aagctaggtc taggtattga 1560

19

tgaagatgaa gtggcagcag aggaacccaa tgctgcagtt cctgatgaga tccccctct 1620
 cgagggcgat gaggatgcgt ctgcgatgcg aggaagtcga gttagggttag gagtcgtact 1680
 tggaacact tgtgcttttg gtttttgtgt ccccatggg gctccactg cgcctcgagt 1740
 gcccc 1745

<210> 22
 <211> 379
 <212> DNA
 <213> Homo sapien

<400> 22
 gtgccaggt ccgcaggtcc gcctggccga gcaggaggcg ccatcatggg agtggacatc 60
 cgccataaca aggaccgaaa ggttcggcgc aaggagccca agagccagga tatctacctg 120
 aggctgttgg tcaagttata caggtttctg gccagaagaa ccaactccac attcaaccag 180
 gttgtgttga agaggttgtt tatgagtcgc accaaccggc cgcctctgtc cttttcccg 240
 atgatccgga agatgaagct tcctggccgg gaaaacaaga cggccgtggg tgtggggacc 300
 ataactgatg atgtgcgggt tcaggaggta cctcgccgcg accacgcac ccatcactg 360
 cgccgctcga catgcatct 379

<210> 23
 <211> 1577
 <212> DNA
 <213> Homo sapien

<400> 23
 ttcccgcggg aatctggacc gcgaccggtt ccacggacgg actccggggc gtccgtggga 60
 gctgggtgcc tgaggacctt ggcggtcggg ccggggcagg agggcgagg tggctgtgat 120
 tcgggatgca cagtaatctg gcggtcagct gcgggcccga ctgggatccg gggcttcggg 180
 ggcgcgcgcc gtccggggtc cgagctggga tcttctgtcg ccgctcacgt actcacttct 240
 gctagtgatg tctggagcta cggaattgtc atgtgggagg tcatgagcta tggagagcga 300
 ccctactggg acatgagcaa ccaggatgtc atcaatgccg tggagcagga ttaccggctg 360
 ccaccacca tggactgtcc cacagactg caccagctca tgctggactg ctgggtgcgg 420
 gaccggaacc tcaggcccaa attctcccag attgtcaata ccctggacaa gctcatccgc 480
 aatgctgcca gcctcaaggt cattgccagc gtcagttctg gcatgtcaca gcccctcctg 540
 gaccgcacgg tcccagatta cacaacctc acgacagttg gtgattggct ggatgccatc 600
 aagatggggc ggtacaagga gagcttcgtc agtcgggggt ttgcatcttt tgacctggtg 660
 gccagatga cggcagaaga cctgctccgt attgggggtc ccctggccgg ccaccagaag 720
 aagatcctga gcagtatcca ggacatgcgg ctgcagatga accagacgct gcctgtgcag 780

gtctgacacc ggctcccacg gggaccctga ggaccgtgca gggatgcaa gcagccggct 840
 ggactttcgg actcttgagc ttttgatgc ctggccttag gctgtggccc agaagctgga 900
 agtttgggaa aggcccaagc tgggacttct ccaggcctgt gttccctccc caggaagtgc 960
 gccccaaacc tcttcatatt gaagatggat taggagaggg ggtgatgacc cctccccaag 1020
 cccctcaggg ccagacactt cctgctctcc agcaggggat cccacaacc tcacacttgt 1080
 ctgttcttca gtgctggagg tcctggcagg gtcaggctgg ggtaagccgg ggttccacag 1140
 ggcccagccc tggcaggggt ctggccccc aggtaggcgg agagcagtcc ctccctcagg 1200
 aactggagga ggggactcca ggaatgggga aatgtgacac caccatcctg aagccagctt 1260
 gcacctccag tttgcacagg gatttgttct gggggctgag ggcctgtcc ccacccccgc 1320
 ccttgggtgct gtcataaaag ggcaggcagg ggcaggctga ggagttgccc tttgcccccc 1380
 agagactgac tctcagagcc agagatggga tgtgtgagtg tgtgtgtgtg tgtgtgctgcg 1440
 cgcgcgcgcg tgtgtgtgtg cacgcactgg cctgcacaga gagcatgggt gagcgtgtaa 1500
 aagcttggcc ctgtgcccta caatggggcc agctgggccg acagcagaat aaaggcaata 1560
 agatgaaaaa aaaaaaa 1577

<210> 24
 <211> 1833
 <212> DNA
 <213> Homo sapien

<400> 24
 tgcaaggcca ggaagctaca gggaaagttc tgatcaaaat acacaaagac acaagccagg 60
 tccccaccgc gcttggcgat gcattcatag cagccttggg gctgtggaca ctccctgggg 120
 cccagcgaag gggagagttt gctcccaaag ggcgaccaat gaccaacatt tgccccccgg 180
 aggaaagaac tggaaccagc ctctgacctg tccagggtgcc ctgtccagct gactgcaagg 240
 acagagagga gtctgcccga gctcttggat cagtctgctg gccgaggagc ccggtggagc 300
 caggggtgac cctggagccc agcctgcccc gaggaggccc cggctcagag ccatgccagg 360
 tgtctgtgat agggccctg acttcctctc cccgtctgaa gaccagggtc tgaggcctgc 420
 cttgggcagc tcagtggctc tgaactgcac ggcttgggta gtctctgggc cccactgctc 480
 cctgccttca gtccagtggc tgaagacgg gcttccattg ggaattgggg gccactacag 540
 cctccacgag tactcctggg tcaaggccaa cctgtcagag gtgcttgtgt ccagtgtcct 600
 gggggtaaac gtgaccagca ctgaagtcta tggggccttc acctgctcca tccagaacat 660
 cagcttctcc tccttctctc ttacagagac tggccctaca agccacgtgg ctgcggtgct 720
 ggctccctc ctggtcctgc tggccctgct gctggccgcc ctgctctatg tcaagtgccg 780

```

tctcaacgtg ctgctctggt accaggacgc gtatggggag gtggagataa acgacgggaa      840
gctctacgac gcctacgtct cctacagcga ctgccccgag gaccgcaagt tcgtgaactt      900
catcctaaag ccgcagctgg agcggcgctcg gggctacaag ctcttcctgg acgaccgcca      960
cctcctgccg cgcgctgagc cctccgccga cctcttggtg aacctgagcc gctgccgacg     1020
cctcatcgtg gtgctttcgg acgccttcct gagccggggc tggtgagacc acagcttcgg     1080
gtgggtcccc cgcggggttg ggtggggccc agcgtacacc caccctctg acggtccccg     1140
cccgagggga gggcctgtgc cggctgctgg agctcaccgc cagaccatc ttcacacct      1200
tcgagggcca gaggcgcgac ccgcgcgacc cggcgctccg cctgctgcgc cagcaccgcc     1260
acctggtgac cttgctgctc tggaggcccg gctccgtgac tccttcctcc gatttttggg     1320
aagaagtgca gctggcgctg ccgcggaagg tgcggtacag gccggtggaa ggagaccccc     1380
agacgcagct gcaggacgac aaggacccca tgctgattct tcgaggccga gtccctgagg     1440
gccggggcct ggactcagag gtggaccgcg accctgaggg cgacctgggt gtccgggggc     1500
ctgttttttg agagccatca gctccaccgc acaccagtgg ggtctcgctg ggagagagcc     1560
ggagcagcga agtggacgtc tcggatctcg gctcgcgaaa ctacagtgcc cgcacagact     1620
tctactgcct ggtgtccaag gatgatatgt agtccacc ccagagtgca ggatcatagg     1680
gacagcgggg gccagggcag cggcgctcgt cctctgctca acaggaccac aaccctgcc      1740
agcagccctg ggacctgcc agcagccctg ggaaaaggct gtggcctcag ggcgcctccc     1800
agtgccagaa aataaagtcc ttttggtatc tga                                     1833

```

```

<210> 25
<211> 2138
<212> DNA
<213> Homo sapien

```

```

<400> 25
tgcaaggcca ggaagctaca gggaaagtgc tgatcaaaat acacaaagac acaagccagg      60
tccccaccgc gcttggcgat gcatccatag cagccttggg gctgtggaca ctccctgggg     120
cccagcgaag gggagagttt gctcccaaag gcgcaccaat gaccaacatt tgccccccgg     180
aggaaagaac tggaaccagc ctctgacctg tccagggtgcc ctgtccagct gactgcaagg     240
acagagagga gtctgcccga gctcttggtg cagtctgctg gccgaggagc ccggtggagc     300
caggggtgac cctggagccc agcctgcccc gaggaggccc cggctcagag ccatgccagg     360
tgtctgtgat agggcccctg acttcctctc cccgtctgaa gaccaggtgc tgaggcctgc     420
cttgggcagc tcagtggctc tgaactgcac ggcttgggta gtctctgggc cccactgctc     480
cctgccttca gtccagtggc tgaaagacgg gcttccattg ggaattgggg gccactacag     540

```

```

cctccacgag tactcctggg tcaaggccaa cctgtcagag gtgcttgtgt ccagtgtcct    600
gggggtcaac gtgaccagca ctgaagtcta tggggccttc acctgtctca tccagaacat    660
cagcttctcc tccttcactc ttcagagagc tggccctaca agccacgtgg ctgcggtgct    720
ggcctccctc ctggctctgc tggccctgct gctggccgcc ctgctctatg tcaagtgccg    780
tctcaacgtg ctgctctggt accaggacgc gtatggggag gtggagataa acgacgggaa    840
gctctacgac gcctacgtct cctacagcga ctgccccgag gaccgcaagt tcgtgaactt    900
catcctaaag ccgcagctgg agcggcgctc gggctacaag ctcttcctgg acgaccgcga    960
cctcctgccg cgcgtgagc cctccgccga cctcttggtg aacctgagcc gctgccgacg   1020
cctcatcgct gtgctttcgg acgccttcct gagccgggcc tggcgcagcc acagcttccg   1080
gtgggtcccg cgcgggggtt ggtgggcccc agcgtacacc cccccctg acggtcccgc   1140
ccgcagggga gggcctgtgc cggctgctgg agctacccg cagacccatc ttcacacct   1200
tcgagggcca gaggcgcgac ccgcgcacc cggcgctccg cctgctgcgc cagcacgcc   1260
acctggtgac cttgctgctc tggaggcccc gctccgtggt gcggagcagg cgcgggaggg   1320
tccggggcta gcggcgggtt agagatgggc ggtgccccgg ctccaggctg ggacccctcc   1380
gtggggagct ctgcggcacc acgctttgtg aatgggccct ggggggaggt tccgctgcct   1440
ggggccccga tgcggggagc cgccttgag gccccggag ccacggaata gctgtcgcag   1500
ggcgtggaac ccgtgggcag ccgcaggtgt gctcttgggg gccaggacgc caggggcttc   1560
cgaggtgttc acacctgcaa accgccccga cctggccccc aggactcctt cctccgattt   1620
ttggaagaaa gtgcagctgg cgctgccgcg gaaggtgcgg tacaggccgg tggaggaga   1680
ccccagacg cagctgcagg acgacaagga cccatgctg attcttcgag gccgagtcct   1740
tgagggcccg gccctggact cagaggtgga cccggaccct gagggcgacc tgggtgtccg   1800
ggggcctggt tttggagagc catcagctcc accgcacacc agtggggtct cgctgggaga   1860
gagccggagc agcgaagtgg acgtctcgga tctcggtcgc cgaaactaca gtgcccgcac   1920
agacttctac tgcttgggtg ccaaggatga tatgtagctc ccaccccaga gtgcaggatc   1980
ataggacag cgggggcccag ggacggcg tcgctcctct gctcaacagg accacaacct   2040
ctgccagcag ccctgggacc ctgccagcag ccctgggaaa aggctgtggc ctcagggcgc   2100
ctccagtgcc cagaaaataa agtccttttg gattctga                               2138

```

```

<210> 26
<211> 676
<212> DNA
<213> Homo sapien

```

23

<400> 26
 cccccgcccc tgcgcccgcg acgcccgcgc tcccgcagag cctccgcgcg acctcttggt 60
 gaacctgagc cgctgccgac gcctcatcgt ggtgctttcg gacgccttcc tgagccgggc 120
 ctggtgcagc cacagcttcc gactccttcc tccgattttt ggaaagaagt gcagctggcg 180
 ctgccgcgga aggtgcggtg caggccggtg gaaggagacc ccagacgca gctgcaggac 240
 gacaaggacc ccatgctgat tcttcgaggc cgagtccttg agggccgggc cctggactca 300
 gaggtggacc cggaccctga gggcgacctg ggtgtccggg ggctgtttt tggagagcca 360
 tcagctccac cgcacaccag tggggtctcg ctgggagaga gccggagcag cgaagtggac 420
 gtctcggatc tcggctcgcg aaactacagt gccgcacag acttctactg cctggtgtcc 480
 aaggatgata tgtagctccc accccagagt gcaggatcat agggacagcg ggggccaggg 540
 cagcggcgtc gctcctctgc tcaacaggac cacaaccctt gccagcagcc ctgggacctt 600
 gccagcagcc ctgggaaaag gctgtggcct cagggcgcct ccagtgcca gaaaataaag 660
 tccttttga ttctga 676

<210> 27
 <211> 1333
 <212> DNA
 <213> Homo sapien

<400> 27
 agatgctgtc gagcggcgca gtgtgatgga tcgtggtcgc ggcgaggaat ggagagcacg 60
 gtctgaatct gcacagagca agatgctgag tggagtcggg ggctttgtgc tgggcctgct 120
 ctcccttggg gccgggctgt tcatctactt caggaatcag aaaggacact ctggacttca 180
 gccaacagga ttcctgagct gaaatgcaga tgaccacatt caaggaagaa ccttctgccc 240
 cagctttgca ggatgaaaag ctttcctgct tggcagttat tcttccacaa gagagggctt 300
 tctcaggacc tggttgctac tggttcggca actgcagaaa atgtcctccc ttgtggcttc 360
 ctgagctcct gcccttggcc tgaagtccca gcattgatgg cagcgcctca tcttcaactt 420
 ttgtgtcccc ctttgccctaa accgtatggc ctcccgtgca tctgtactca ccctgtacga 480
 caaacacatt acattattaa atgtttctca aagatggagt taaatatcat ctggtccatt 540
 tggctccaaa gacagaaaat gaaaagaaaa agggaagatt attttcctaa tagaataatg 600
 attttcatgt atatgtcatg agtgtgtgag gtaatgcgta tgttaaatag ctggatttag 660
 acattccaca ctatagccat atatcaaac ttcattgctgt acaatataaa tatmctatac 720
 aatttttact tgtcaattaa aaaagtaatc ctaacattta aaaaggcaat gcataaaaac 780
 tgagaacaga ctataacaac tgaaacaaac tgggcaaaca tgagatgaga aaccagccag 840
 caagtcaatc agaactcttt cttcatctcg tctacaatat tttgtattta taactgtaaa 900

ttagtgata	gtgtttcact	ccagagactt	caataatata	gtgttatcaa	aggacttgta	960
cagatttcag	agaaagacac	attdagaaga	tggagagttc	tccgttatgt	tctatctgag	1020
agtcagtatg	aaatgtcaaa	tctaaaagta	cataattcag	aggcttatct	caaattaata	1080
atcatttgag	cataatttct	ccactgtcag	agacgactgt	tattttatct	tcaatcaaat	1140
taaaagtgtg	ttttatgcat	atcttatttt	tagttatatg	ttacttgtag	atacgtagca	1200
gtacaagtgc	atataaatcc	tgtaagagta	taatcctata	gaaatatatt	aagctgataa	1260
ttatgtctgt	tctgtttgat	cacagagttg	caacaaatag	gcctcatttc	ctaagttgag	1320
gaaatgattt	tct					1333

<210> 28
 <211> 1228
 <212> DNA
 <213> Homo sapien

<400> 28	
ctttgggagc	tggtctcttc cataagcttt tctggaggag gaacwrtggc ttgctgagg 60
acactctgga	cttcagccaa caggattcct gagctgaaat gcagatgacc acattcaagg 120
aagaaccttc	tgccccagct ttgcaggatg aaaagctttc ctgcttggca gttattcttc 180
cacaagagag	ggctttctca ggacctggtt gctactgggt cggcaactgc agaaaatgtc 240
ctcccttggt	gcttcctcag ctccctgcct tggcctgaag tcccagcatt gatggcagcg 300
cctcatcttc	aacttttggt ctcccttttg cctaaaccgt atggcctccc gtgcatctgt 360
actcaccctg	tacgacaaac acattacatt attaaatggt tctcaaagat ggagttaaat 420
atcatctggt	ccatttggtt ccaaagacag aaaatgaaaa gaaaaaggga agattatttt 480
cctaatagaa	taatgatttt catgtatatg tcatgagtgt gtgaggtaat gcgtatgtta 540
aatagctgga	tttagacatt ccacactata gccatatatc aaaacttcat gctgtacaat 600
ataaatatmc	tatacaattt ttacttgtag attaaaaaag taatcctaac atttaaaaag 660
gcaatgcata	aaaactgaga acagactata acaactgaaa caaactgggc aaacatgaga 720
tgagaaaacca	gccagcaagt caatcagaac tctttcttca tctcgtctac aatattttgt 780
atttataact	gtaaattagt gtatagtgtt tcaactccaga gacttcaata atacagtgtt 840
atcaaaggac	ttgtacagat ttcagagaaa gacacattta gaagatggag agttctccgt 900
tatgttctat	ctgagagtca gtatgaaatg tcaaacttaa aagtacataa ttcagaggtc 960
tatttcaaat	taataatcat ttgagcataa tttctccact gtcagagacg actgttattt 1020
tattttcaat	caaattaaaa gttgttttta tgcatatctt atttttagtt atatgttact 1080
tgtagacacg	tagcagtaca agtgcataa aatcctgtaa gagtataatc ctatagaaat 1140

atattaagct gataattatg tctgttctgt ttgatcacag agttgcaaca aataggcctc 1200
atttcctaag ttgaggaaat gattttct 1228

<210> 29
<211> 2619
<212> DNA
<213> Homo sapien

<400> 29
ggggaacaca acttaggcta ggtacgaggc ctgggtgggc mggctatcgg actctccatt 60
gccctgcacg tgggttttga caacttcgag cagctgctct cgggggctca ctggatggac 120
cagcacttcc gcacgacgcc cctggagaag aacgcccccg tcttgctggc cctgctgggt 180
atctggtaca tcaactgctt tgggtgtgag acacacgcca tgctgcccta tgaccagtac 240
ctgcaccgct ttgctgcgta cttccagcag ggcgacatgg agtccaatgg gaaatacatc 300
accaaactcg gaaccctgtg ggaccaccag acaggcccca ttgtgtgggg ggagccaggg 360
accaatggcc agcatgcttt ttaccagctc atccaccaag gcaccaagat gataccctgt 420
gacttctca tcccggcca gaccagcac ccatacggg agggctctgca tcacaagatc 480
ctcctggcca acttcttggc ccagacagag gccctgatga ggggaaaatc gacggaggag 540
gcccgaagg agctccaggc tgcgggcaag agtccagagg accttgagag gctgctgcca 600
cataaggtct ttgaaggaaa tcgccaacc aactctattg tgttcaccaa gctcacacca 660
ttcatgcttg gagccttggc cgccatgtat gagcacaaga tcttcgttca gggcatcatc 720
tgggacatca acagctttga ccagtgggga gtggagctgg gaaagcagct ggctaagaaa 780
atagagcctg agcttgatgg cagtgtcaa gtgacctctc acgacgcttc taccaatggg 840
ctcatcaact tcatcaagca gcagcgcgag gccagagtc aataaactcg tgctcatctg 900
cagcctcctc tgtgactccc ctttctcttc tcgtccctcc tccccggagc cggcactgca 960
tgttcttga caccaccag agcaccctct ggttgtgggc ttggaccacg agcccttagc 1020
agggaaggct ggtctcccc agcctaacct ccagccctc catgtctatg ctccctctgt 1080
gttagaattg gctgaagtgt ttttgtgcag ctgacttttc tgacctatgt tcacgttggt 1140
cacatcccat gtagaaaaat aaagatgcca cggaggagggt ttagggtca gcctctgatt 1200
tttttttcc tgtgatggtg ctttatgtag cagagggcag gagcgctcag caggacgcag 1260
gctgtgcctc tgcggacact taacactaag tggtagcggt gtctagagtg gagcaagggt 1320
ccctgagaag acaatagtgg ggtgggggca caatcagtca ggacggcaca gktrcraggct 1380
cayrcctgta atcycagcac tttgggaggc craggyaggk ggatcacctg aggtcagacc 1440
agtctggtca acatggtgaa acccgtctg taccaaaaat acaaaaacca gctgggcatg 1500

```

gtggcacctg cctgtggtcc cagctactcg ggaggccgag gcagaagact cgctggaacc 1560
cgggaggcgg agcttgccgt gagccgagac tgtgtcactg caccgcagcc tgggtgacag 1620
agccagactc catctcaaca acagacgcac ttacacttgg agtgtctgtc gcacagggcc 1680
gaggagcgca tgtcacacag gcggtggtc cctttgctac tgctgtagac gaacacgttg 1740
cactggtgcg ggtggaactc ggctgcagtg atgacttcgg tcagctcctc catgttagca 1800
ggcttgatgt ccacgatgtt aaagcttcta tctgtgattt ctaagtgcc aataattaatt 1860
ctcaggtcat ctgcagaaag atatgtttca tgatcactat ttactgaaat ggaatttata 1920
tgatatgtgt gagcatttgc aaaaattcgc cgtggactcg cttctaccat aagatccatg 1980
ggcttcaata ttgggacccg tagcgccgtg atcctaaatg ggtctcgaag tcttccatct 2040
tcgtctttca gggtataacc ttctgctctt ttatcccgtt cacttatttt ccataatttt 2100
atagttttat catttgtaga cagtagaaaa tgagcagcat tctgttgttg taaccaccta 2160
attttattaa ttttttctc aatttctaga cttttcaa atgtcaaac cggttcatga 2220
ctttgaaagg tgctgtaaac attatattct ccctagaat gagggcggtt tttattctct 2280
tgttcacgct gaaaaataac aactctgccg cccttgctc ctgttgcaag aagatctcca 2340
gagtaattaa actcaacggt ggaaatgatg tccgcttcgg ccacgtcctc gtcgatggcc 2400
cccttgacct gcgagaagca ccactggaag tcgttgccgc ccgcggggca gccgccgctt 2460
ccggctcctg ccatggcagc cggcgccggg acggcgggg accgcggcag accagggaga 2520
cggcctccgc gccgctgcag ccgagcgctc agccccggg ccgccaccac cgccggcgcc 2580
gccgccgccc cgccgccgc cggggagggg tyttyyket 2619

```

```

<210> 30
<211> 2564
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (554)..(554)
<223> n=a, c, g or t

```

```

<400> 30
actggtgtgt agttcttaaa gagcagttat cactgttccc tgccttgttg tgtcatcgga 60
tgtgtccccc tcgctccaac tgagctgtac agcgagtgc ccaggctggg actgaccctt 120
gctgagaagt accaggcggg cttgtccctc tgctgaagc ccagacagt ctctcaagca 180
taactgatgt ctgctcatc agatcctcct ggccaacttc ttggcccaga cagaggccct 240
gatgagggga aaatcgacgg aggaggcccg aaaggagctc caggctgcgg gcaagagtcc 300

```


agaggacctt gagaggctgc tgccacataa ggtcagcact tctgcatttg gctttggggt	360
gcatgctgga gttggagggtg tgaagctatg gcaccaggca ggggcttggg gcatgcctgg	420
cttgtcctac agaatgggag ggctgtcct cacagacctg cacgtctcag cctctggggc	480
aggggtgtgct tcccttcaag gggtttaggg atcaggactc tcttggagac attccttgggt	540
gttttctgca ggtntttgaa ggaaatcgcc caaccaactc tattgtgttc accaagctca	600
caccattcat gcttgagacc ttggtcgcca tgtatgagca caagatcttc gtccagggca	660
tcatctggga catcaacagc tttgaccagt ggggagtgga gctgggaaag cagctggcta	720
agaaaataga gcctgagctt gatggcagtg ctcaagtgac ctctcacgac gcttctacca	780
atgggctcat caacttcac aagcagcagc gcgaggccag agtccaataa actcgtgctc	840
atctgcagcc tcctctgtga ctcctcttc tcttctcgtc cctcctccc ggagccggca	900
ctgcatgttc ctggacacca cccagagcac cctctggttg tgggcttggga ccacgagccc	960
ttagcaggga aggctggtct cccccagcct aacccccagc cctccatgt ctatgctccc	1020
tctgtgttag aattggctga agtggttttg tgcagctgac ttttctgacc catgttcacg	1080
ttgttcacat cccatgtaga aaaataaaga tgccacggag gaggtttag gctcagcctc	1140
tgattttttt tttcctgtga tgggtgcttta tgtagcagag ggcaggagcg ctacgagga	1200
cgcaggctgt gcctctgcgg aacttaaca ctaagtgggtg agcgggtcta gagtggagca	1260
aggtgccctg agaagacaat agtgggggtg gggcacaatc agtcaggacg gcacagktrc	1320
rggctcayrc ctgtaatcyc agcactttgg gaggccragg yaggkggatc acctgaggtc	1380
agaccagtct ggtcaacatg gtgaaacccc gtctgtacca aaaatacaaa aaccagctgg	1440
gcatggtggc acctgcctgt ggtcccagct actcgggagg ccgaggcaga agactcgctg	1500
gaaccgggga ggcggagctt gccgtgagcc gagactgtgt cactgcaccg cagcctgggt	1560
gacagagcca gactccatct caacaacaga cgcacttaca cttggagtgt ctgtcgaca	1620
gggccgagga gcgcatgtca cacaggcgga tggcccttt gctactgctg tagacgaaca	1680
cgttgcactg gtgcgggtgg aactcggtg cagtgatgac ttcggtcagc tcctccatgt	1740
tagcaggctt gatgtccacg atgttaaagc ttctatctgt gatcttctaa tgccataaat	1800
taattctcag gtcactctga gaaagatatg tttcatgac actatttact gaaatggaat	1860
ttatatgata tgtgtgagca ttgcaaaaa ttcgccgtgg actcgcttct accataagat	1920
ccatgggctt caatattggg acccgtagcg ccgtgaccc aaatgggtct cgaagtcttc	1980
catcttcgtc tttcagggtta taacctctctg ctcttttctc ccgttcactt attttccata	2040
attttatagt tttatcattt gtagacagta gaaaatgagc agcattctgt tgtggtaacc	2100

28

acctaatttt attaattttt tcttcaattt ctagactttt caaatagtca aactccggtt 2160
 catgactttg aaaggtgctg taaacattat atttctccctt agaatgaggg cggtttttat 2220
 tctcttggtc acgtgaaaa ataacaactc tgccgccctt gtctcctgtt gcaagaagat 2280
 ctccagagta attaaactca acggtggaaa tgatgtccgc ttcggccacg tctcgtcga 2340
 tggcccccctt gacctgcgag aagcaccact ggaagtcgtt gccgcccgcg gggcagccgc 2400
 cgctccggc tcttgcctg gcagccggcg gcgggacggg cggggaccgc ggcagaccag 2460
 ggagacggcc tccgcgccgc tgcagccgag cgctcagccc cggggccgcc accaccgccg 2520
 gcgccgccgc cgccgccgcc gccgccgggg agggatytty ykct 2564

<210> 31
 <211> 5718
 <212> DNA
 <213> Homo sapien

<400> 31
 catcaatgaa aactgtgctc agcattaacc ttcagaagct gatgacactg ttctacttta 60
 agcaggaaag agaagttttc tccgcacctg ttgagggcac gtcttctctc gcgcctgttg 120
 agggcacttc tctaaagtga atggatcctg ctttttgccg acataggtgc gtggaattgc 180
 cagcgctctc cagaacatca agactgcact ggggagaacc cacagaaacg gacaggactg 240
 ctggccaggt ggcttcccc gcaggcgccc cagagctagg tgacgacacc tccagggaac 300
 gtgcgccgca gacctcaagg cagtggcagt tccacgggag cagcagtgcc acgtggaacc 360
 tctacacacg ggaccgcctt gggctgcacc ttctgctctg ccctgggggt ggaggagggtg 420
 ggctatttag tgggaaatac attctagagg aagatggaaa aggggtgtcca aatacgcagt 480
 gaaaatccag ttaaacattt aaaatttatt gaactttttg gtcaaaatag ttgaacaaat 540
 atatacaatc ccattttact agaactgtct taaggaccag ggttgactaa tgcttaata 600
 actgttccac acctcagact ttgtaaattg tgtcacaatc aatatacagt atataaatta 660
 ttttttagtt aaaataaagg tacctttaga ttctcaattt ctaacttaat gtgaactcaa 720
 gtggcactgt gttagcattg attcagggta atgacggaca tctgttctt gcccatgtgg 780
 attagtaatc agcacacacc ttcattctcca cagggggatg ggggatcagc acacaccttc 840
 atctccacag ggggaccggg gatcagcaca cgcttcatc tccacagggg gatgggggat 900
 cagcacacac cttcatctcc acagtgggat ggcggatcag cacacgcctt catctccaca 960
 gggggatggg ggatgagcac acgccttcat ctccacaggg ggatggcgga tcagcacacg 1020
 ccttcatctc cacaggggga cggcggatca gcacacgcct tcatctccac agggggacgg 1080
 gggatcagca cagccttca tctccacagg aggacggggg atcagcacac accttcatct 1140

ccacaggggg accgggggatc agcacacacc ttcatctcca caggggggatg ggggatcagc 1200
acacaccttc atctccacag ggggatgggg gatcagcaca cgccttcac tccacagggg 1260
gatgggggat cagcacacgc cttcatctcc acagggggat gggggatcag cacacgcctt 1320
catctccaca gtgggatggc ggatcagcac acgccttcac ctccacaggg ggatggggga 1380
tcagcacaca cttcatctc cacaggggga tgggggatca gcacacgcct tcatctccac 1440
agtgggatgg cggatcagca cagccttca tctccacag gggatggggg atcagcacac 1500
gccttcacat ccacagtggg acgggggatc agcacacacc tccatctcca cagggggacg 1560
ggggaccagc acacacctcc atctccacag gggcacgagg aagcagcaca caccttcac 1620
tccacaggag gacgggggat cagcacacat cttcatctcc acagggggat gggggatcag 1680
cacacacctc catctccaca gggggacaag agatcagcac acaccttcac cccacaggg 1740
ggacagggga tcagcatacc cttcatctc cacaggggga tgggggatca gcacacacct 1800
tcatctccac agggggacgg gggatcagca tacccttca tctccacag gggatggggg 1860
atcagcacac accttcaccc ccacagggg atgggggatc agcacacacc tccatctcca 1920
caggggcatg aggaatcagc acacacctc atctccacag gaagacgggg gatcagcaca 1980
ccccttcac tccacagggg gatgggggat cagcacacac cttcatctcc acagggggac 2040
gggggatcag cagtacacac cttcatctag ataggggaac aggggatcag tacacacctc 2100
tatctccaca ggggcacgag ggatcagcac acacctccat ctccacaggg catgggggat 2160
cagcacaccc cttcatctcc acagggggac ggggatcagc acacacctc atctccacag 2220
ggggacgggg gatcaacaca ccccttcac tccacatggg gatgggggat cagcacacac 2280
cttcatctcc acagggggac gggggatcag cacacgcctc catgtccaca ggggcacggg 2340
ggatcagcac accccttcac cttcacaggg agatggggga tcagcacaca cttcatctc 2400
cacaggggga cggggatcag cacacacat ctccacaggg ggacagggga tcaacacacc 2460
ccttcacctc cacatgggga tgggggatca gcacacacct tcatctccac aggggggatg 2520
ggatcagcac acgcctccat ctccacaggg gcgcggggga ttggcacacc cttcatctc 2580
cacagggaga tgggggatca gcacacacct tcatctccac agggggacgg gggatcagca 2640
cacccttca tctccacagt gggacagggg atcagcatat cccttcacat ccacaggggg 2700
atggggatca gcacacacct tcatctccac agggagatgg gggatcagca cgcaccttca 2760
tctccacag ggcacggggg atcagcacac accttcacat ccacaggggg acgggggatc 2820
agcacacacc ttcatccaga caggggcatg aggaatcagc acacacctc atctagacag 2880
ggggatggga tgggggatca gcacacacct ccatctccac aggggcacga gggagcagca 2940
cacacctcca tctccacag gcacagggga tcagcacaca cctccttctc cacaggggca 3000

caggcaatca gcacatacct tcatctagac aggggtgacgg gcgatcagca cacacctcca	3060
tctccacagc agcaggagga atcaacacac accttcatct ccacaggcgg acgsggggatc	3120
agcacacacc tccatctcca caggggcacg ggagcccgac tccccactcc cctcggggac	3180
acgtagtggg cgtgtgttaa atccacagga agcaagctct ggtcactcag catttgtaca	3240
gctctgtatt atttgtcttg gaaaatcttg atttctggta tcaaagatcc tactttgaaa	3300
ataaaccaac gtactggtgg catcgtgtaa ttaaaattgc ttaattttgc tggcttgaaa	3360
tgcaggatga aatttagcat gcatgtaatt ttagcaatgt ttaaaaagta tactttcccc	3420
aaagtgcttt ttcaaaagcc ttagaaagca agcacgtatt tcttcaactg tagtgaattt	3480
gaagctactt taaaccaggc ttactggatg gcagcaattt tggttcttcc ttgcaatttt	3540
ctttctcaga ggtggcctcg gttctggctt tcccagacc ttacctggtg tatctgatgg	3600
gctacttcta cagaaacaca gcgggcagca gcttggggac agaaatgctg cacacatacc	3660
cccacgctgc caaagaagga cacacatttc tcaaacagtt agttccaaag aagaaccctg	3720
cataacttca ttctatgcaa taaaattaaa taaaaagaca atacttctgt ttggatttaa	3780
aaacacttta tagtgtgttg tttttattga gcgctcacac ccgaaggggt ggcgggcgat	3840
gctgtgggag tggagctggg aaagcagctg gctaagaaaa tagagcctga gcttgatggc	3900
agtgtcaag tgacctctca cgacgttct accaatgggc tcatcaactt catcaagcag	3960
cagcgcgagg ccagagtcca ataaactcgt gctcatctgc agcctcctct gtgactcccc	4020
tttctcttct cgtccctcct ccccgagcc ggactgcat gttcctggac accaccaga	4080
gcaccctctg gttgtgggct tggaccacga gcccttagca gggaggctg gtctcccca	4140
gcctaacccc cagccctcc atgtctatgc tccctctgtg ttagaattgg ctgaagtgtt	4200
tttgtgcagc tgacttttct gacctgtt cacgttggtc acatcccatg tagaaaaata	4260
aagatgccac ggaggagggt gtaggctcag cctctgattt ttttttctct gtgatggtgc	4320
tttatgtagc agagggcagg agcgtcagc aggacgcagg ctgtgcctct gcggacactt	4380
aacactaagt ggtgagcggg tctagagtgg agcaagggtg cctgagaaga caatagtggg	4440
gtgggggcac aatcagtcag gacggcacag ktrcrggctc ayrcctgtaa tcycagcact	4500
ttgggaggcc raggyaggkg gatcacctga ggtcagacca gtctgggtcaa catggtgaaa	4560
ccccgtctgt accaaaaata caaaaaccag ctgggcatgg tggcacctgc ctgtggctcc	4620
agctactcgg gaggccgagg cagaagactc gctggaaccc gggaggcgga gcttgccgtg	4680
agccgagact gtgtcactgc accgcagcct ggggtgacaga gccagactcc atctcaacaa	4740
cagacgcact tacacttgga gtgtctgtcg cacagggccg aggagcgcat gtcacacagg	4800

31

cggatggtcc ctttgctact gctgtagacg aacacgttgc actggtgcgg gtggaactcg 4860
 gctgcagtga tgacttcggt cagctcctcc atgttagcag gcttgatgtc cacgatgtta 4920
 aagcttctat ctgtgatttc taagtgccat aaattaattc tcaggtcacg tgcagaaaga 4980
 tatgtttcat gatcactatt tactgaaatg gaatttatat gatatgtgtg agcatttgca 5040
 aaaattcgcc gtggactcgc ttctaccata agatccatgg gcttcaatat tgggacccgt 5100
 agcgccgtga tcctaaatgg gtctcgaagt cttccatctt cgtctttcag gttataacct 5160
 tctgctcttt tatcccgctc acttattttc cataatttta tagttttatc attttagac 5220
 agtagaaaaat gagcagcatt ctgttggtgt aaccacctaa ttttattaat ttttctctca 5280
 atttctagac ttttcaaata gtcaaaactcc gggtcatgac tttgaaagggt gctgtaaaca 5340
 ttatattctc ccctagaatg agggcggtt ttattctctt gttcacgctg aaaaataaca 5400
 actctgccgc ccttgctctc tgttgcaaga agatctccag agtaattaaa ctcaacgggtg 5460
 gaaatgatgt ccgcttcggc cacgtcctcg tcgatggccc ccttgacctg cgagaagcac 5520
 cactggaagt cggtgccgcc cgcggggcag ccgcccctc cggctcctgc catggcagcc 5580
 ggcggcgggg cgggcgggga ccgcccagga ccaggagagc ggcctccgcg ccgctgcagc 5640
 cgagcgctca gccccggggc cgccaccacc gccggcgccg ccgcccgcgc cgccgcccgc 5700
 ggggagggat yttykct 5718

<210> 32
 <211> 1091
 <212> DNA
 <213> Homo sapien

<400> 32
 gcctttgcgs gtggcggcga acgaggagag cacgccatga aggcctcggg cacgctacga 60
 gagtacaagg tagtgggtcg ctgcctgcc acccccaaat gccacacgcc gcccctctac 120
 cgcattcgaa tctttgcgcc taatcatgtc gtcgccaagt cccgcttctg gtactttgta 180
 tctcagttaa agaagatgaa gaagtcttca ggggagattg tctactgtgg gcagggtgtt 240
 gagaagtccc ccctgcgggt gaagaacttc gggatctggc tgcgctatga ctcccggagc 300
 ggcaccaca acatgtaccg ggaataccgg gacctacca ccgaggcgc tgtcaccag 360
 tgctaccgag acatgggtgc ccggcaccgc gcccgagccc actccattca gatcatgaag 420
 gtggaggaga tcgcgccag caagtgccgc cggccggctg tcaagcagtt ccacgactcc 480
 aagatcaagt tcccgtgcc ccaccgggtc ctgcgccgtc agcacaagcc acgcttcacc 540
 accaagaggc ccaacaacct tctttctagg tgcagggcc tcgtccgggg tgtgccccca 600
 aataaactca ggaacgcccc gaaggtaaaa gctgctgttc cggatgactt taaggacttc 660

32

agtctgctgt aacgaggagg cgcgctaata tcagttcaag accatggtga ggcgagcttg 720
 gagcgctggc acgcatgacc cggagaagag caccggaaac cgggatgggg aacggctgga 780
 tgccaagtca agcgccagaa gatgggccaa gcgggaccgg accacacgac gagcactgcc 840
 cgcttaagaa gaataccaca gcaacgcgaa agcaacagta aggcagaaca aaccgaggcg 900
 acaccagagc ggcgcgaaag aaaaaaacgc agcacaacga acacgcggca gcgcaatacg 960
 cagcgcgctc taaagagacc gacagaaagc aaccgcgcgg ggacaaccaa ggcgagacga 1020
 agccccggg acgcaaacgg gagggagaag agcgcaccgc aggtccgaat aaagagagga 1080
 acagccgaca t 1091

<210> 33
 <211> 859
 <212> DNA
 <213> Homo sapien

<400> 33
 acagccgtgt tgtcgccagg ccgcgccttc cctccacag cgcgcgctgc gcgtgcgaag 60
 gtctggcggc tcttgggact ggcggggctg cgcgcggggt tagggtaggg gtacgggaag 120
 gctcaaccca ggacctgctt accttgcttt gggggcgac taagcacctg ccgggagcag 180
 ggggcgcacc gggaactcgc agatttcgcc agttgggcgc actggggatc tgtggactgc 240
 gtccggggga tgggctaggg ggacatgcgc acgctttggg ccttacagaa tgtgatcgcg 300
 cgagggggag ggcgaagcgt ggcgggaggg cgaggcgaag gaaggagggc ctcgcttcgtc 360
 cggcccagcc tgaccaatgc cgcggtggaa acgggcttgg agctggcccc ataagggctg 420
 gcggcttctt ccgaagccgc cctcccccac agcttctcga ctgcagtggg gcggggggca 480
 ccaacacttg gagatttttc cggaggggag aggatcttct aagggcacag agaattcatt 540
 ttctacacat taacttgagc tgctggaggg aactgctgg caaacggaga cctatttttg 600
 taaaaaagac ccttgacctg ggcgtaata aagatgacct ggaccctgc cccactatc 660
 tggagtcttc catgctggcc aagatctgga cagagcagt ccctgagggg cggggtccct 720
 ggcgtgaggc ccccgtaga gccaccctg gggtaggggt gtgggactg ctgctctgct 780
 agggagaagc ctgtgtggg cacacctctt caagggagcg tgaactttat aaataaatca 840
 gttctgttta ccaaaaaa 859

<210> 34
 <211> 2236
 <212> DNA
 <213> Homo sapien

<400> 34
 ctcacccta agcccagccc tcggctccca cccttggtga cctgggccga accattcacc 60

ggagcgcgca gcgggtggag tgtggctcgg aggaccgcgg cgggtcaagc acctttctcc 120
cccatactctg aaagcatgcc ctttgtccac gtctgttacg ctcatataaa cttccagaat 180
gcaacaggac ggacttggag tagggacaag gaacggaagt ggggaagggga ggagcgtgca 240
cccctcctgg ccttgggtcg cgccgcgccc cctaaggtac tttggaaggg acgcgcgggc 300
cagacgcgcc cagacggccg cgatggcgct gttggccggc gggctctcca gagggctggg 360
ctcccacccg gccgccgcag gccgggacgc ggtcgtcttc gtgtggcttc tgcttagcac 420
ctggtgcaca gtccttgcca ggccatcca ggtgaccgtg tccaaccctt accacgtggt 480
gatectcttc cagcctgtga ccctgccctg tacctaccag atgacctga cccccacgca 540
accatcgtc atctggaagt acaagtcttt ctgccgggac cgcatcgccg atgccttctc 600
ccgggccagc gtcgacaacc agctcaatgc ccagctggca gccgggaacc caggctacaa 660
cccctacgtc gagtgccagg acagcgtgcg caccgtcagg gtctgtggca ccaagcaggg 720
caacgctgtg accctgggag attactacca gggccggagg attaccatca ccggaatgc 780
tgacctgacc tttgaccaga cggcgtgggg ggacagtggg gtgtattact gtcctgtggt 840
ctcagcccag gacctccagg ggaacaatga ggccacgca gagctcatcg tccttgggag 900
gacctcaggg gtggctgagc tcttacctgg ttttcaggcg gggcccatag aagactggct 960
cttcgtggtt gtggtatgcc tggtgcctt cctcatcttc ctctctctgg gcatctgctg 1020
gtgccagtgc tgcccgcaca cttgctgctg ctacgtcagg tgcccctgct gccagacaa 1080
gtgctgctgc cccgaggccc tgtatgccgc cggcaaagca gccacctcag gtgttcccag 1140
catttatgcc cccagacct atgcccacct gtctcccgcc aagacccac ccccaccagc 1200
tatgattccc atgggccctg cctacaacgg gtaccctgga ggataccctg gagacgttga 1260
caggagtagc tcagctgggt gccaaaggct ctatgtacct ctgcttcggg acacggacag 1320
cagtgtggcc tctgaagtcc gcagtggcta caggattcag gccagccagc aggacgactc 1380
catgcgggtc ctgtactaca tggaagaagg gctggccaac ttcgaccctt ctgcacctgg 1440
cccccccagt ggccgtgtgg agcggggccat gagtgaagtc acctccctcc acgaggacga 1500
ctggcgatct cggccttccc ggggccctgc cctaccccg atccgggatg aggagtgggg 1560
tgccactcc ccccgagtc ccaggggatg ggaccaggag cccgccaggg agcaggcagg 1620
cgggggctgg cgggccaggc gggcccgggc ccgctccgtg gacgccctgg acgacctcac 1680
cccgccgagc accgccgagt caggagcag gtctccacg agtaatggtg ggaggagaag 1740
ccgggcctac atgccccgcg ggagccgcag ccgggacgac ctctatgacc aagacgactc 1800
gagggacttc ccacgtccc gggaccccca ctacgacgac ttcaggtctc gggagcgccc 1860

34

tcctgccgac cccaggtccc accaccaccg taccgaggac cctcgggaca acggtccag 1920
 gtccggggac ctcccctatg atgggaggct actggaggag gctgtgagga agaaggggtc 1980
 ggaggagagg aggagacccc acaaggagga ggaggaagag gcctactacc cggccgcgcc 2040
 gccccgtac tcggagaccg actcgcaggc gtcccgagag cgcagggtca agaagaactt 2100
 ggcctgagt cgggaaagt tagtcgtctg atctgacgtt ttctacgtag cttttgtatt 2160
 ttttttttta atttgaagga aactgatga agcctgcca taccctccc gagtctaata 2220
 aaacgtataa tcacaa 2236

<210> 35
 <211> 1612
 <212> DNA
 <213> Homo sapien

<400> 35
 ggataacttga tttcggcttg tggggacagt ggtggacca gcatctgggc tttatataaa 60
 gggcagcttt gttgccctgt aaacacacag accatgggtg gccacttctt ccagtaagtt 120
 agctggggag ttggaagttt aggtaaaacc ttttgattga caaatgttg cgaattacca 180
 tgctgttaaa tgaaacattg ttctgccacc ctggggctgt ggggtgcctgc gtgcaccctc 240
 tgaaaaatca cacaggaagt ggggtggggt ctctgtgaag ctgggtgtccc ccagcctcag 300
 ggatgtctga gaaatggaat gaggaccaac agggactcag atgtccaagg aagctctaca 360
 gcggagagga cggcttggga aggaggtcca ggcccagggt cctccggaac ccaatgggta 420
 tggggcagcc tggctcctgc ctcatcccc ttctcctgtt gattgtgtcc tcacagtgt 480
 tgccgccggc aaagcagcca ctcaggtgt tccagcatt tatgccccca gcacctatgc 540
 ccacctgtct cccgccaaga cccaccccc accagctatg attcccatgg gccctgccta 600
 caacgggtac cctggaggat accctggaga cgttgacagg agtagctcag ctgggtggcca 660
 aggtccttat gtaccctgc ttcgggacac ggacagcagt gtggcctctg aagtccgcag 720
 tggctacagg attcaggcca gccagcagga cgactccatg cgggtcctgt actacatgga 780
 gaaggagctg gccaaactcg acccttctcg acctggcccc ccagtgggcc gtgtggagcg 840
 ggccatgagt gaagtcacct ccctccacga ggacgactgg cgatctcggc cttcccgggg 900
 ccctgccctc accccgatcc gggatgagga gtgggggtggc cactcccccc ggagtcccag 960
 gggatgggac caggagcccg ccaggagga ggcaggcggg ggctggcggg ccaggcggcc 1020
 ccggggccgc tccgtggacg ccctggacga cctaccccc ccgagcaccg ccgagtcagg 1080
 gagcaggtct cccacgagta atgggtgggag gagaagccgg gcctacatgc cccgcgggag 1140
 ccgcagccgg gacgacctct atgaccaaga cgactcgagg gacttcccac gctcccggga 1200

35

```

ccccactac gacgacttca ggtctcggga gcgccctcct gccgacccca ggtcccacca 1260
ccaccgtacc cgggaccctc gggacaacgg ctccagggtcc ggggacctcc cctatgatgg 1320
gcggctactg gaggaggctg tgaggaagaa ggggtcggag gagaggagga gacccccaaa 1380
ggaggaggag gaagaggcct actacccgcc cgcgccgcc cgtactcgg agaccgactc 1440
gcaggcgtcc cgagagcgca ggctcaagaa gaacttggcc ctgagtcggg aaagttagt 1500
cgtctgatct gacgttttct acgtagcttt tgtatttttt tttttaattt gaaggaacac 1560
tgatgaagcc ctgccatacc cctcccgagt ctaataaaac gtataatcac aa 1612

```

```

<210> 36
<211> 6003
<212> DNA
<213> Homo sapien

```

```

<400> 36
tgacatacga aagggttttta atatctacaa aagatgacag actgaaatag taacattctc 60
caggatggaa agatctatgc cacgtaggaa ttgtttgttc aaattattat tttttaaaag 120
ccataagcta tggagagaat aacagagttt tgttttgttt tttaatcctg atacggtgga 180
agtctggtag gatgggaacc gttcttgaag cttgagaata cattcggaac aaataaaaag 240
taggaaatgc atacattttg ttattttacag ttgtagggca ggtaagggca aaacgtattg 300
ataaactcaa agaattttta gaatgatatt gtcattgagcc atgagccatg gaagctgcaa 360
ggtttgggct ccctcctgcc atcgggtggg ggggtgtcggc ggaagggggg acgtgtcctg 420
tcctctccca gtgcttattt cctgctggaa ccgtggtttt gacgcagccg tgacagccct 480
tcctccttat taaaatcaca caatgaagca ggtcgcccaa gattgaaagc gcgtttccag 540
ggcggacccc ggtccccgcc tttggccagc aaaagtacat tctggcaggg agggatctct 600
gagccccgaa gtgagggtcat gccagttagc agttagaaaa ggaggcattt ctacacttg 660
ctaaagagtt tgacgcttgt gaaatataaa caggaagcgg gagctctcca aggcgacctg 720
aaggaaatac actgccagtt tgggtcaact gtgaatgcca gttctctagg tcgtcaaggg 780
cactgccagt taagactgat ccctaactg cccccgaggc tactactctt cgcccagttc 840
cggcagtgac ccattttttt tttcaatcct tcccacagcc aggtcactca gcagtggggt 900
gactgacccg aagcgcagtg ggaagggtt acagtcatct ctcaactccc cgaacttgca 960
gggcagggga gcgcagaggt cagtgaatc cccagccagg ccttatttta gtcattgagt 1020
tccactggct ccctcctctg actgtatcag cgctgtcac accgcgggat tggagtttca 1080
gctgttttcc ctgctcccg ctgagtgaga aggtcacaga gcccgacct ttccggggga 1140
ggtgggtggg gccctccggc agcctaagt agattgcaga gatctggggc cgctggagga 1200

```

gggtccttccg ccccgggctt ggggctttgg gttagcttct cccccgcttc tctacctagg 1260
 catctcggtt acaatgaaat taaaaacaaa acacaaagca acacaagttc tcattgtttt 1320
 tccccctgtg cagccagtgg cctcgataca attttcccag ggctgccttt cctttattgc 1380
 tcaatttaac ctcttcctag aggttctgca agttcaaggg caggatccgt gtcggattca 1440
 tcaactagtg cagcgcttag tgcagagcag ggggctcaag aaatacgtgt cgaccgcatg 1500
 attgctaaga tttccctgta acaggcggtt tttttttctt gtctctctca tccctcctgc 1560
 cctgggtgca gaccgcgaac agcgccgccg aacttcccaa caggcaggga gagggccgat 1620
 ccgggctggc acgccacgtc cccggggtct cagtccctgg gtcaacgcag agaggctcgc 1680
 gggtccgcgtt gctacgacct gagcccggcc acgtgaagcg tcacgccctg cgctgctgca 1740
 gggcaggcgc cgccgccgcg gttgtgtaac gctgcagccg gaggggaggc gaccggggcg 1800
 cctctgctcc ggagggggca gaggaggag ggagcgccgg gcacagcgcc cggtatcttt 1860
 ctgcgcttga ctggccggga ggaaggggtt gacactgggg caccagggg gctgcgcaac 1920
 cgggtgcccg ccgcccgcga gggtgtgccc aggagcggag ctcttgggag cgagggggccc 1980
 tcgggtcaca ccctccacgg tcaggtgcgc gggcggtgct cggcggcagc ctgagctgcc 2040
 cctgcgcttg gccgtcggcc cgttcccat ggggtgacat ccgccccgcc cctcggtccc 2100
 tccccaggc gggcaattcc tggacgcgag ggtgagcagt gggggcagg aagccaggac 2160
 ggaaagaaac ccagcctct ggggaaggct ggggggccgg gcgactccct ccgcagcgcg 2220
 ccggttcctc caagcgggcc gggcggggga ggaggaagag ggctgggctg gagctagcaa 2280
 ggggatattc ctctcccggc ttgagtcaga cgcggcgga tccgtcctcc cccgttccct 2340
 cccaggagac gggaacttac ttcatttccc tggggcaggt tcgcccacgt taccaacttc 2400
 tccccctccc ccagcaccct cgtcccttcc agcttccgcg cccccaccc aactgggcag 2460
 gaccaggtc gtgctgccac cccctcttcg gggaaaggcg gccgcagccg cagacacctg 2520
 ggggccgggg ctgggggtgg gggctcccta gcagccgccg gagcgttgtc caacacgtga 2580
 gactcatgtg atgaagccgg gggagggcgg gcaggtcgct ccttccctcc ccggcagtgg 2640
 ccagacgtgc ctggagtcac agggtagaac acgtagctcc aaccaccca ccgctccctc 2700
 ccgctcccc gcccgccca cttctcattc acttggtcgc cagggcgag acagaccgcg 2760
 cagggagcac acaccgccag tctgtgcgct gagtcggagc cagaggccgc ggggacaccg 2820
 ggccatgcac gcccccaact gaagctgcat ctcaaagccg aagattccag cagcccaggg 2880
 gatttcaaag agctcagact cagaggaaca tctgcggaga gacccccgaa gccctctcca 2940
 gggcagtcct catccagacg ctccgctagt gcagacagga gcgcgcagtg gccccggctc 3000
 gccgcgcat ggagcggatc ccagcgcgc aaccaccccc cgctgcctg cccaaagcac 3060

cgggactgga gcacggagac ctaccagggg tgtaccctgc ccacatgtac caagtgtaca	3120
agtcaagacg ggggaataaag cggagcgagg acagcaaggt aagcaagtgc acccctaggg	3180
accctgcgct cagcccctcg cgcgcgctga gtttccaaga aaagttttct cgctttgagg	3240
ttggcgaggg gatgcagggt gtgggctcgg tgcctcttct gagtgacttg gaaaagaaa	3300
ggcagacgat ggttctgggt gctactctgc tactctgctc ctctgccgga ctgttgctgc	3360
ggggctggga ggataggctt ctcatctcct tccccaagcg cccaccgcct ccccgtcgt	3420
cttgacggag acctacaaat tgccgcaccg gctcatcgag aaaaagagac gtgaccggat	3480
taacgagtgc atcgcccagc tgaaggatct cctaccgaa catctcaaac ttacaacttt	3540
gggtcacttg gaaaaagcag tggttcttga acttaccttg aagcatgtga aagcactaac	3600
aaacctaatt gatcagcagc agcagaaaat cattgccctg cagagtgggt tacaagctgg	3660
tgagctgtca gggagaaatg tcgaaacagg tcaagagatg ttctgctcag gtttccagac	3720
atgtgcccgg gaggtgcttc agtatctggc caagcacgag aacactcggg acctgaagtc	3780
ttcgcagctt gtcaccacc tccaccgggt ggtctcggag ctgctgcagg gtggtacctc	3840
caggaagcca tcagaccag ctcccaaagt gatggacttc aaggaaaaac ccagctctcc	3900
ggccaaagggt tcggaaggtc ctgggaaaaa ctgctgcca gtcattccagc ggactttcgc	3960
tcactcgagt ggggagcaga gcggcagcga cacggacaca gacagtgggt atggaggaga	4020
atcggagaag ggcgacttgc gcagtgagca gccgtgcttc aaaagtgacc acggacgcag	4080
gttcacgatg ggagaaagga tcggcgcaat taagcaagag tccgaagaac cccccacaaa	4140
aaagaaccgg atgcagcttt cggaatgatga aggccatttc actagcagtg acctgatcag	4200
ctccccgttc ctgggcccac acccacacca gcctccttct tgcttgcct tctacctgat	4260
cccaccttca gcgactgcct acctgcccac gctggagaag tgctggatc ccacctcagt	4320
gccagtgcta taccagggc tcaacgcctc tgccgcagcc ctctctagct tcatgaaccc	4380
agacaagatc tcggctccct tgctcatgcc ccagagactc ccttctccct tgccagctca	4440
tccgtccgtc gactcttctg tcttgctcca agctctgaag ccaatcccc ctttaaactt	4500
agaaacaaa gactaaactc tctaggggat cctgctgctt tgctttcctt cctcgctact	4560
tcctaaaaag caacaaaaaa gtttttgtga atgctgcaag attgttgcat tgtgtatact	4620
gagataatct gaggcattga gacgagattc aggggtgtgtg tgtgtgtgtg tgtgtgtgtg	4680
tgtatgtgcg tgtgctgca catgtgtgcc tgcgtgttg tataggactt taaagctcct	4740
tttggcatag ggaagtcacg aaggattgct tgacatcagg agacttgggg gggattgtag	4800
cagacgtctg ggcttttccc caccagaga atagccccct tcgatacaca tcagctggat	4860

38

ttctaaaagc ttcaaagtct tggctctgtga gtcactcttc agtttgggag ctgggtctgt 4920
 ggctttgatc agaaggtact ttcaaaagag ggctttccag ggctcagctc ccaaccagct 4980
 gttaggaccc cacccttttg cctttattgt cgacgtgact caccagacgt cggggagaga 5040
 gagcagtcag accgagcttt ctgctaacat ggggaggttag caggcactgg catagcacgg 5100
 tagtggtttg gggaggtttc cgcaggtctg ctccccaccc ctgcctcgga agaataaaga 5160
 gaatgtagtt ccctactcag gctttcgtag tgattagctt actaaggaac tgaaaatggg 5220
 ccccttgtag aagctgagct gccccggagg gagggaggag ttccctgggc ttctggcacc 5280
 tgtttctagg cctaaccatt agtacttact gtgcaggga ccaaaccaag gtctgagaaa 5340
 tgcggacacc ccgagcgagc accccaaagt gcacaaagct gagtaaaaag ctgccccctt 5400
 caaacagaac tagactcagt ttcaattcc atcctaaaac tccttttaac caagcttagc 5460
 ttctcaaagg cctaaccaag ccttggcacc gccagatcct ttctgtaggc taattcctct 5520
 tgcccaacgg catatggagt gtccttattg ctaaaaagga ttccgtctcc ttcaaagaag 5580
 ttttattttt ggtccagagt acttgttttc ccgatgtgtc cagccagctc cgcagcagct 5640
 ttctaaaatg cactatgcct gattgctgat cgtgttttaa ctttttcttt tcctgttttt 5700
 attttggtat taagtcgttg cctttatttg taaagctgtt ataaatatat attatataaa 5760
 tatattaaaa aggaaaatgt ttcagatgtt tatttgtata attacttgat tcacacagtg 5820
 agaaaaaatg aatgtattcc tgtttttgaa gagaagaata attttttttt tctctagggg 5880
 gaggtacagt gtttatattt tggagccttc ctgaaggtgt aaaattgtaa atatttttat 5940
 ctatgagtaa atgttaagta gttgttttaa aatacctaata aaaataattc ttttcctgtg 6000
 gaa 6003

<210> 37
 <211> 5871
 <212> DNA
 <213> Homo sapien

<400> 37
 tgacatacga aagggttttta atatctacaa aagatgacag actgaaatag taacattctc 60
 caggatggaa agatctatgc cacgtaggaa ttgtttgttc aaattattat tttttaaag 120
 ccataagcta tggagagaat aacagagttt tgttttgttt tttaatcctg atacgggtgga 180
 agtctggtag gatgggaacc gttcttgaag cttgagaata cattcggaac aaataaaaag 240
 taggaaatgc atacattttg ttatttacag ttgtagggca ggtaagggca aaacgtattg 300
 ataaactcaa agaatttttaaat gaatgatatt gtcatgagcc atgagccatg gaagctgcaa 360
 gggttgggct ccctcctgcc atcgggtggg ggggtgtcggc ggaagggggg acgtgtcctg 420

39

tcctctccca gtgcttattt cctgctggaa ccgtgggttt gacgcagccg tgacagccct	480
tcatccttat taaaatcaca caatgaagca ggtcgccc aa gattgaaagc gcgtttccag	540
ggcggacccc ggtccccgcc tttggccagc aaaagtacat tctggcaggg agggatctct	600
gagccccgaa gtgaggtcat gccagttagc agttagaaaa ggaggcattt ctcacacttg	660
ctaaagagtt tgacgcttgt gaaatataaa caggaagcgg gagctctcca aggcgacctg	720
aaggaaatac actgccagtt tggtaactt gtgaatgcca gttctctagg tcgtcaaggg	780
cactgccagt taagactgat cccctaactg caccgagggc tactactctt cgcccagtct	840
cggcagtgc cccatttttt tttcaatcct tcccacagcc aggtcactca gcagtggggt	900
gactgacccg aagcgcagtg ggaagggctt acagtcactt ctcaactccc cgaacttgca	960
gggcagggga gcgcagaggt cagtgaatc cccagccagg ccttatttta gtcagtgtt	1020
tccactggct ccctcctctg actgtatcag cgctgctcac accgcgggat tggagtttca	1080
gctgttttcc ctgctcccg ctgagtgaga aggtcacaga gcccggaacct ttccggggga	1140
gggtgggtggg gccctccggc agcctaagtg agattgcaga gatctggggc cgctggagga	1200
ggtccttccg ccccgggctt ggggctttgg gttagcttct ccccgcttc tctacctagg	1260
catctcggtt acaatgaaat taaaaacaaa acacaaagca acacaagttc tcattgtttt	1320
tccctctgt cagccagtgg cctcgataca attttcccag ggctgccttt cctttattgc	1380
tcaatttaac ctcttcttag aggttctgca agttcaaggg caggatccgt gtcggattca	1440
tcaactagtg cagcgcctag tgcagagcag ggggctcaag aaatacgtgt cgaccgcatg	1500
attgctaaga ttccctgta acaggcggtt ttttttctt gtctctctca tccctcctgc	1560
cctgggtgca gaccgcgaac agcgcgcggc aacttcccaa caggcagggg gagggccgat	1620
ccgggctggc acgccacgtc cccggggctt cagtccttg gtcaacgcag agaggctcgc	1680
ggtcgcgggt gctacgacct gagcccggc acgtgaagcg tcacgccctg cgctgctgca	1740
gggcaggcgc cgccgcgcg gttgtgtaac gctgcagccg gaggggaggg gaccggggcg	1800
cctctgctcc ggagggggca gaggagggag ggagcgcgg gcacagcgcc cggtatcttt	1860
ctgcgcttga ctggccggga ggaaggggt gacactgggg caccagggg gctgcgcaac	1920
cggtgcccc cccgcccga ggttgtgccc aggagcggag ctctgggag cgagggggcc	1980
tcggctcaca ccctccacgg tcaggtgctc gcggcgtgct cggcggcagc ctgagctgcc	2040
cctgcgcttg gccgtcggc cgcttcccat ggggtgacat cggccccgc cctcggtccc	2100
tccccaaggc gggcaattcc tggacgcgag ggtgagcagt gggggcaggg aagccaggac	2160
ggaaagaaac ccagcctct ggggaaggct ggggggcccg gcgactccct ccgcagcgcg	2220
ccggttcctc caagcgggccc gggcggggga ggaggaagag ggctgggctg gagctagcaa	2280

ggggatattc ctctcccggc ttgagtcaga cgcgggcgga tccgtcctcc cccgttcctc	2340
cccaggagac gggaacttac ttcatattccc tggggcaggt tcgcccacgt taccaacttc	2400
tccccctccc ccagcacccc cgtcccttcc agcttcgcg cccccaccc aactgggcag	2460
gacccaggtc gtgctgccac cccctcttcg gggaaaggcg gccgcagccg cagacacctg	2520
ggggccgggg ctgggggtgg gggctcccta gcagccgccg gagcgttgtc caacacgtga	2580
gactcatgtg atgaagccgg gggagggcgg gcaggtcgct ccttccctcc ccggcagtgg	2640
ccagacgtgc ctggagtcac agggtagaac acgtagctcc aaccaccca ccgctccctc	2700
ccgcctcccc gcccgcacca cttctcattc acctggctcg cacggcgag acagaccgcg	2760
cagggagcac acaccgccag tctgtgcgct gagtcggagc cagaggccgc ggggacaccg	2820
ggccatgcac gcccctaact gaagctgcat ctcaaagccg aagattccag cagcccaggg	2880
gatttcaaag agctcagact cagaggaaca tctgcggaga gacccccgaa gccctctcca	2940
gggcagtcct catccagacg ctccgctagt gcagacagga gcgcgcagtg gcccggctc	3000
gccgcgccat ggagcggatc cccagcgcgc aaccaccccc cgcctgcctg cccaaagcac	3060
cgggactgga gcacggagac ctaccaggga tgtacctgc ccacatgtac caagtgtaca	3120
agtcaagacg gggaataaag cggagcgagg acagcaaggt aagcaagtgc acccctaggg	3180
accctgcgct cagcccctcg cgcgcgctga gttccaaga aaagttttct cgctttgagg	3240
ataggcttct catctcttc cccaagcgc caccgcctcc ccgtgcgtct tgcaggagac	3300
ctacaaattg ccgcaccggc tcatcgagaa aaagagacgt gaccggatta acgagtgc	3360
cgcccagctg aaggatctcc taccgaaca tctcaaactt acaactttgg gtcacttgga	3420
aaaagcagtg gttcttgaac ttaccttgaa gcatgtgaaa gcactaacia acctaattga	3480
tcagcagcag cagaaaatca ttgccctgca gagtggttta caagctgggtg agctgtcagg	3540
gagaaatgtc gaaacaggtc aagagatgtt ctgctcaggt ttccagacat gtgcccggga	3600
ggtgcttcag tatctggcca agcacgagaa cactcgggac ctgaagtctt cgcagcttgt	3660
cacccacctc caccgggtgg tctcggagct gctgcagggg ggtacctcca ggaagccatc	3720
agaccagct cccaaagtga tggacttcaa ggaaaaacc agctctccgg ccaaagggtc	3780
ggaaggctct gggaaaaact gcgtgccagt catccagcgg actttcgctc actcgagtgg	3840
ggagcagagc ggacgcgaca cggacacaga cagtggctat ggaggagaat cggagaaggg	3900
cgacttgccg agtgagcagc cgtgcttcaa aagtgaccac ggacgcaggt tcacgatggg	3960
agaaaggatc ggcgcaatta agcaagagtc cgaagaaccc cccacaaaaa agaaccggat	4020
gcagctttcg gatgatgaag gccatttcac tagcagtgac ctgatcagct ccccgctcct	4080

gggcccacac ccacaccagc ctcctttctg cctgcccttc tacctgatcc caccttcagc 4140
gactgcctac ctgcccattg tggagaagtg ctggtatccc acctcagtgc cagtgtctata 4200
cccaggcctc aacgcctctg ccgcagccct ctctagcttc atgaaccag acaagatctc 4260
ggctcccttg ctcatgcccc agagactccc ttctcccttg ccagctcatc cgtccgtcga 4320
ctctttctgtc ttgtcccaag ctctgaagcc aatccccctt ttaaacttag aaaccaaaga 4380
ctaaactctc taggggatcc tgctgctttg ctttccttcc tcgctacttc ctaaaaagca 4440
acaaaaaagt ttttgtgaat gctgcaagat tgttgcatg tgtatactga gataatctga 4500
ggcatggaga gcagattcag ggtgtgtgtg tgtgtgtgtg tgtgtgtgtg tatgtgcgtg 4560
tgcgtgcaca tgtgtgcctg cgtgttggtg taggacttta aagctccttt tggcataggg 4620
aagtcacgaa ggattgcttg acatcaggag acttgggggg gattgtagca gacgtctggg 4680
cttttcccca ccagagaaat agcccccttc gatacacatc agctggattt tcaaaagctt 4740
caaagtcttg gtctgtgagt cactcttcag tttgggagct gggctctgtg ctttgatcag 4800
aagggtacttt caaaagaggg ctttccaggg ctccagctccc aaccagctgt taggacccca 4860
cccttttgcc tttattgtcg acgtgactca ccagacgtcg gggagagaga gcagtcagac 4920
cgagctttct gctaacatgg ggaggtagca ggcactggca tagcacggta gtggtttggg 4980
gaggtttccg caggtctgct cccacccctt gcctcggag aataaagaga atgtagttcc 5040
ctactcaggc tttcgtagt attagcttac taaggaaactg aaaatgggcc ctttgtacaa 5100
gctgagctgc ccggagggga gggaggagtt ccctgggctt ctggcacctg tttctaggcc 5160
taaccattag tacttactgt gcagggaacc aaaccaaggt ctgagaaatg cggacacccc 5220
gagcgagcac cccaaagtgc acaaagctga gtaaaaagct gcccccttca aacagaacta 5280
gactcagttt tcaattccat cctaaaactc cttttaacca agcttagctt ctcaaaggcc 5340
taaccaagcc ttggcacgc cagatccttt ctgtaggcta attcctcttg cccaacggca 5400
tatggagtgt ccttattgct aaaaaggatt ccgtctcctt caaagaagtt ttatttttgg 5460
tccagagtac ttgttttccc gatgtgtcca gccagctccg cagcagcttt tcaaaatgca 5520
ctatgcctga ttgctgatcg tgttttaact ttttcttttc ctgtttttat tttggatta 5580
agtcgttgcc tttatttgta aagctgttat aaatatatat tatataaata tattaaaaag 5640
gaaaatgttt cagatgttta tttgtataat tacttgatc acacagtgag aaaaaatgaa 5700
tgtattcctg tttttgaaga gaagaataat ttttttttcc tctagggaga ggtacagtgt 5760
ttatatattg gagccttct gaaggtgtaa aattgtaaatt atttttatct atgagtaaat 5820
gttaagtagt tgttttaaaa tacttaataa aataattctt ttcctgtgga a 5871

<210> 38
<211> 7742
<212> DNA
<213> Homo sapien

<400> 38
tgacatacga aagggttttta atatctacaa aagatgacag actgaaatag taacattctc 60
caggatggaa agatctatgc cacgtaggaa ttgtttgttc aaattattat tttttaaaag 120
ccataagcta tggagagaat aacagagttt tgttttgttt tttaatcctg atacgggtgga 180
agtctggtag gatgggaacc gttcttgaag cttgagaata cattcggaac aaataaaaag 240
taggaaatgc atacattttg ttatttacag ttgtagggca ggtaagggca aaacgtattg 300
ataaactcaa agaattttta gaatgatatt gtcatgagcc atgagccatg gaagctgcaa 360
ggtttgggct ccctcctgcc atcgggtggg ggggtgtcggc ggaagggggg acgtgtcctg 420
tcctctccca gtgcttattt cctgctggaa ccgtggtttt gacgcagccg tgacagccct 480
tcctccttat taaaatcaca caatgaagca ggtcgcccaa gattgaaagc gcgtttccag 540
ggcggacccc ggtccccgcc tttggccagc aaaagtacat tctggcaggg agggatctct 600
gagccccgaa gtgaggtcat gccagttagc agttagaaaa ggaggcattt ctcacacttg 660
ctaaagagtt tgacgcttgt gaaatataaa caggaagcgg gagctctcca aggcgacctg 720
aaggaaatac actgccagtt tgggtcaact gtgaatgcca gttctctagg tcgtcaaggg 780
cactgccagt taagactgat cccctaactg caccgaggc tactactctt cgtccagtct 840
cggcagtgc cccatttttt tttcaatcct tcccacagcc aggtcactca gcagtggggg 900
gactgacccg aagcgcagtg ggaagggtt acagtcatct ctcaactccc cgaacttgca 960
gggcagggga gcgcagaggt cagtgaatc cccagccagg ccttatttta gtcatgagtt 1020
tcactgggt ccctcctctg actgtatcag cgctgtcac accgcgggat tggagtttca 1080
gctgttttcc ctgctcccg ctgagtgaag aggtcacaga gcccggacct ttccggggga 1140
gggtgggtggg gccctccggc agcctaagt agattgcaga gatctggggc cgctggagga 1200
ggtccttccg ccccggtt ggggttttg gttagcttct ccccgttc tctacctagg 1260
catctcggtt acaatgaaat taaaaacaaa acacaaagca acacaagttc tcattgtttt 1320
tccctctgt cagccagtgg cctcgatata attttccag ggctgcctt cctttattgc 1380
tcaatttaac ctcttcttag aggttctgca agttcaaggg caggatccgt gtcggattca 1440
tcaactagt cagcgcctag tgcagagcag ggggtcaag aaatacgtgt cgaccgcatg 1500
attgctaaga tttccctgta acaggcggtt ttttttctt gtctctctca tccctcctgc 1560
cctgggtgca gaccgcgaac agcgcgccg aacttcccaa caggcagga gagggccgat 1620
ccgggtggc acgccagtc cccgggtct cagtccttg gtcaacgcag agaggctcgc 1680

gggtccgcggt gctacgacct gagccccggcc acgtgaagcg tcacgcccctg cgctgctgca	1740
gggcaggcgc cgccgcccgc gtgtgtgtaac gctgcagccg gaggggagggc gaccgggggcg	1800
cctctgctcc ggagggggca gaggagggag ggagcgcccg gcacagcgcc cggtatcttt	1860
ctgcgcttga ctggccggga ggaaggggggt gacactgggg caccagggg gctgcgcaac	1920
cggtgccccg cccgcccga ggtgtgtccc aggagcgag ctcctgggag cgagggggccc	1980
tcggctcaca ccctccacgg tcagggtgcgc gcggcgctgct cggcggcagc ctgagctgcc	2040
cctgcgcttg gccgtcggcc cgcttcccat ggggtgacat ccgccccgc cctcggtccc	2100
tccccaaggc gggcaattcc tggacgcgag ggtgagcagt gggggcaggg aagccaggac	2160
ggaaagaaac ccagccctct ggggaaggct ggggggcccgc gcgactccct ccgcagcgcg	2220
ccggttcctc caagcgggcc gggcggggga ggaggaagag ggctgggctg gagctagcaa	2280
ggggatattc ctctcccgcc ttgagtcaga cgcgggcgga tccgtcctcc cccgttccct	2340
cccaggagac gggaacttac ttcatttccc tggggcaggt tcgcccacgt taccaacttc	2400
tccccctccc ccagcacccc cgtcccttcc agcttccgcg cccccaccc aactgggcag	2460
gaccaggtc gtgctgccac cccctcttcg gggaaaggcg gccgcagccg cagacacctg	2520
ggggccgggg ctgggggtgg gggctcccta gcagccgccc gagcgttgtc caacacgtga	2580
gactcatgtg atgaagcccg gggagggcg gcaggtcgct ccttccctcc ccggcagtg	2640
ccagacgtgc ctggagtcac agggtagaac acgtagctcc aaccaccca ccgctccctc	2700
ccgcctcccc gcccgcccca cttctcattc acttggtcgc cacggcgag acagaccgcg	2760
caggagcac acaccgccag tctgtgcgct gagtgcgagc cagaggccgc ggggacaccg	2820
ggccatgcac gcccacaact gaagctgcat ctcaaagccg aagattccag cagcccagg	2880
gatttcaaag agctcagact cagaggaaca tctgcggaga gacccccgaa gccctctcca	2940
gggcagtcct catccagacg ctccgctagt gcagacagga gcgcgcagtg gcccggctc	3000
gccgcgcat ggagcggatc ccagcgcgc aaccaccccc cgctgcctg cccaaagcac	3060
cgggactgga gcacggagac ctaccagggt aagttggcac tccttggccc ttcaagcctc	3120
aactgcagcc ccatgttatg cgccactctc aacttgagc agctccgggc gcaccgggag	3180
gttcttggcc ggcagaacgc ctgcagccgt gcgcaagcag tcacgacctc tcctgggtctc	3240
tcttctgca ggatgtaccc tgccacatg taccaagtgt acaagtcaag acggggaata	3300
aagcggagcg aggacagcaa ggtaagcaag tgcacccta gggaccctgc gctcagcccc	3360
tcgcgcgcgc tgagtttcca agaaaagttt tctcgctttg aggttggcga ggggatgcag	3420
gggtgtgggt cggtgcctct tctgagtgac ttggaaaaga aagggcagac gatggttctg	3480

ggtgctactc tgctactctg ctctctgcc ggactgttgc tgcggggctg ggaggatagg 3540
 cttctcatct ccttcccaa gcgcccaccg cctccccgtg cgtcttgca gagacctaca 3600
 aattgccgca cgggctcatc gagaaaaaga gacgtgaccg gattaacgag tgcacgccc 3660
 agctgaagga tctcctaccc gaacatctca aacttacagt aagtgagaag ctggccccctt 3720
 tccaaccag tcttttggcc agagggcggg cggttcactc gccgcctgca ggttccgagg 3780
 ggaactgcag accagactgg cggatcagag ctggcggtg gctctgcggt ggggtcctcc 3840
 tccacagtcc tggcgttttc tggctgactc atgcaaata ggcgagaacc acttggaact 3900
 ctccaagtgt gctccaacac caactggcct tctttttcag gagtttaact tctgactttg 3960
 ggtcacttgg aaaaagcagt ggttcttgaa cttaccttga agcatgtgaa agcactaaca 4020
 aacctaattg atcagcagca gcagaaaatc attgccctgc agagtgggtt acaagctggt 4080
 gagtgtgat tctggctatg ctctcttta gagtttgagt gcaaataag agcccgtagg 4140
 ctcaacatct gatgtaataa acattactgc acaaattgc ttatgcacgt tcattggggg 4200
 agatgcagtt ttatctttcc taaaaggta gacttttcat ttgaaatgca ttgtgccacc 4260
 tggcatgaat acgtattagt gaaatgataa gatattcctg tggcatttta gcttgagagt 4320
 ctctgggttc tgggagattt cctactccc ttttaaaatg ccaaggcacc tataatttgc 4380
 attgtgggtg ttttacagtt ggtcttctta gccctcttcc attggagaag aaatcatggg 4440
 tatgccccag cgttttctg tttctttgct tcttgtcac ttggtgtgtg gtgtgatect 4500
 ggcaggatac agcaggagt tttcttgc ttcattgtgt tccctgaata gccactagg 4560
 atcttgggtg attgcacaat ctacgaaac cactgaagt tcatcattcg tttctactga 4620
 cttccagata atcgagtgca accttctaac cttctagtct cacttcttcc aaataatact 4680
 gtacagactg gggagaatta ttctaccac tccctcatt catgcttgc tgccttctc 4740
 tcgaaggctc gtatgatgaa aattgcaaaa acccagctag tgcttctgcc tgtttcctca 4800
 ttcaaaatag gaagatggtc atacctaata tcttgaggat catggataga atatgtggag 4860
 aagacataga aatagccta ttaaaggta cagcatgact gcaacttta caaatcattg 4920
 aaaataagtt tactttagg gattgttgt tcttcattct tttctcatt ttataactat 4980
 ccagcctctg cccattttgc caggtgcaa tgtgtatgc attgactctt aggtttataa 5040
 atacatcaga atgttctttg ctgcctcgt tacatggaaa actcaciaaac cacctataca 5100
 atccaaaaga ggggaaacag ctcatctcat attaattatg gtccatttcc atgataggat 5160
 attatgcaac cattaaaaat catattttat ctgagtgtta gtatatggat tattttaatt 5220
 atttatactt ttcagtgtaa ttcttctctg agcactact ttttgaggt aggaacacaa 5280
 atcagtaaat gcattttttt ttttaacaaa aactaaagca cctgggtttt tggagtatag 5340

tttccaggtg	cgccacaggc	cttctgaaac	gttttccttc	ggatttttct	ttccccaggt	5400
gagctgtcag	ggagaaatgt	cgaaacaggt	caagagatgt	tctgctcagg	tttccagaca	5460
tgtgcccggg	aggtgcttca	gtatctggcc	aagcacgaga	acactcggga	cctgaagtct	5520
tcgcagcttg	tcacccacct	ccaccgggtg	gtctcgagc	tgctgcaggg	tggtagctcc	5580
aggaagccat	cagacccagc	tcccaaagt	atggacttca	aggaaaaacc	cagctctccg	5640
gccaaaggtt	cggaaggtcc	tgggaaaaac	tgctgccag	tcattccagc	gactttcgct	5700
cactcgagt	gggagcagag	cggcagcgac	acggacacag	acagtggcta	tggaggagaa	5760
tcggagaagg	gcgacttg	cagtgcagc	ccgtgcttca	aaagtgacca	cggacgcagg	5820
ttcacgatgg	gagaaaggat	cggcgcaatt	aagcaagagt	ccgaagaacc	ccccacaaaa	5880
aagaaccgga	tgcagcttcc	ggatgatgaa	ggccatttca	ctagcagtga	cctgatcagc	5940
tccccgttcc	tgggcccaca	cccacaccag	cctcctttct	gcctgccctt	ctacctgatc	6000
ccaccttcag	cgactgccta	cctgcccatt	ctggagaagt	gctggatatc	cacctcagt	6060
ccagtgtat	accaggcct	caacgcctct	gccgcagccc	tctctagctt	catgaaccca	6120
gacaagatct	cggctccctt	gctcatgccc	cagagactcc	cttctccctt	gccagctcat	6180
ccgtccgtcg	actcttctgt	cttgctccaa	gctctgaagc	caatcccccc	tttaaactta	6240
gaaaccaaag	actaaactct	ctaggggagc	ctgctgcttt	gctttccttc	ctcgctactt	6300
cctaaaaagc	aacaaaaaag	tttttgtgaa	tgctgcaaga	ttgttgcat	gtgtatactg	6360
agataatctg	aggcatggag	agcagattca	gggtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	6420
gtatgtgcgt	gtgcgtgcac	atgtgtgcct	gcgtgttggt	ataggacttt	aaagctcctt	6480
ttggcatagg	gaagtcacga	aggattgctt	gacatcagga	gacttggggg	ggattgtagc	6540
agacgtctgg	gcttttcccc	accagagaa	tagccccctt	cgatacacat	cagctggatt	6600
ttcaaaagct	tcaaagtctt	ggtctgtgag	tcactcttca	gtttgggagc	tgggtctgtg	6660
gctttgatca	gaaggtactt	tcaaaagagg	gctttccagg	gctcagctcc	caaccagctg	6720
ttaggacccc	acccttttgc	ctttattgtc	gacgtgactc	accagacgtc	ggggagagag	6780
agcagtcaga	cagagcttcc	tgctaacatg	gggaggtagc	aggcactggc	atagcacggt	6840
agtggtttgg	ggagggttcc	gcagggtctg	tccccacccc	tgctcgggaa	gaataaagag	6900
aatgtagttc	cctactcagg	ctttcgtagt	gattagctta	ctaaggaaact	gaaaaatgggc	6960
cccttgatca	agctgagctg	ccccggaggg	agggaggagt	tccctgggct	tctggcacct	7020
gtttctaggc	ctaaccatta	gtacttactg	tgaggggaac	caaaccaagg	tctgagaaat	7080
gcggacaccc	cgagcgagca	ccccaaagt	cacaaagctg	agtaaaaagc	tgcccccttc	7140

46

```

aaacagaact agactcagtt ttcaattcca tcctaaaact ccttttaacc aagcttagct 7200
tctcaaaggc ctaaccaagc cttggcaccg ccagatcctt tctgtaggct aattcctctt 7260
gcccacaggc atatggagtg tccttattgc taaaaaggat tccgtctcct tcaaagaagt 7320
tttatttttg gtccagagta cttgttttcc cgatgtgtcc agccagctcc gcagcagctt 7380
ttcaaaatgc actatgcctg attgctgatc gtgttttaac tttttctttt cctgttttta 7440
ttttggtatt aagtcggtgc ctttatttgt aaagctgtta taaatatata ttatataaat 7500
atattaaaaa ggaaaatggt tcagatgttt atttgtataa ttacttgatt cacacagtga 7560
gaaaaaatga atgtattcct gtttttgaag agaagaataa tttttttttt ctctagggag 7620
aggtacagtg tttatatattt ggagccttcc tgaaggtgta aaattgtaa tattttttatc 7680
tatgagtaaa tgtaagtag ttgttttaaa atacttaata aaataattct tttcctgtgg 7740
aa 7742

```

```

<210> 39
<211> 5384
<212> DNA
<213> Homo sapien

```

```

<400> 39
tgacatacga aagggttttta atatctacaa aagatgacag actgaaatag taacattctc 60
caggatggaa agatctatgc cacgtaggaa ttgtttgttc aaattattat tttttaaaag 120
ccataagcta tggagagaat aacagagttt tgttttgttt tttaatcctg atacgggtgga 180
agtctggtag gatgggaacc gttcttgaag cttgagaata cattcggaac aaataaaaag 240
taggaaatgc atacattttg ttattttacag ttgtagggca ggtaagggca aaacgtattg 300
ataaactcaa agaattttta gaatgatatt gtcatgagcc atgagccatg gaagctgcaa 360
ggtttgggct ccctcctgcc atcgggtggg gggtgtcggc ggaagggggg acgtgtcctg 420
tcctctccca gtgcttattt cctgctggaa ccgtggtttt gacgcagccg tgacagccct 480
tcatccttat taaaatcaca caatgaagca ggtcgcccaa gattgaaagc gcgtttccag 540
ggcggacccc ggtccccgcc tttggccagc aaaagtacat tctggcaggg agggatctct 600
gagccccgaa gtgaggtcat gccagttagc agttagaaaa ggaggcattt ctcacacttg 660
ctaaagagtt tgacgcttgt gaaatataaa caggaagcgg gagctctcca aggcgacctg 720
aaggaaatac actgccagtt tggtaactt gtgaatgcca gttctctagg tcgtcaaggg 780
cactgccagt taagactgat cccctaactg caccgaggc tactactctt cgcccagtct 840
cggcagtgac ccattttttt tttcaatcct tcccacagcc aggtcactca gcagtggggg 900
gactgacccg aagcgagtg ggaagggtt acagtcattc ctcaactccc cgaacttgca 960

```

gggcagggga gcgcagaggt cagtgaatc cccagccagg ccttatTTTA gtcagagtt 1020
tccactggct ccctcctctg actgtatcag cgctgctcac accgcgggat tggagtttca 1080
gctgttttcc ctgctcccgg ctgagtga gaagtcacaga gcccggaacct ttccggggga 1140
gggtgggtggg gccctccggc agcctaagtg agattgcaga gatctggggc cgctggagga 1200
ggtccttccg ccccgggctt ggggctttgg gttagcttct cccccgcttc tctacctagg 1260
catctcgggt acaatgaaat taaaaacaaa acacaaagca acacaagttc tcattgtttt 1320
tccctctgt cagccagtgg cctcgataca attttcccag ggctgccttt cctttattgc 1380
tcaatttaac ctcttcctag aggttctgca agttcaaggg caggatccgt gtcggattca 1440
tcaactagtg cagcgcctag tgcagagcag ggggctcaag aaatacgtgt cgaccgcatg 1500
attgctaaga ttccctgta acaggcgggt ttttttctt gtctctctca tccctcctgc 1560
cctgggtgca gaccgcgaac agcgccgccc aacttccaa caggcagga gagggccgat 1620
ccgggctggc acgccacgtc cccggggtct cagtccctgg gtcaacgcag agaggctcgc 1680
ggtcgcgggt gctacgacct gagcccgcc acgtgaagcg tcacgacctg cgctgctgca 1740
gggcagggcg cgccgcccgg gttgtgtaac gctgcagccg gaggggaggc gaccggggcg 1800
cctctgctcc ggagggggca gaggaggag ggagcgccg gcacagcgcc cggtatcttt 1860
ctgcgcttga ctggccggga ggaagggggg gacactgggg caccagggg gctgcgcaac 1920
cgggtgcccg cccgcccga ggtgtgccc aggagcgag ctctgggag cgagggggccc 1980
tcggctcaca ccctccacgg tcaggtgcgc gggcggtgct cggcggcagc ctgagctgcc 2040
cctgcgcttg gccgtcggcc cgttcccat ggggtgacat ccgccccgc cctcggctcc 2100
tccccaaagg gggcaattcc tggacgcgag ggtgagcagt gggggcagg aagccaggac 2160
ggaaagaaac cccagcctct ggggaaggct ggggggcccg gcgactccct ccgcagcgcg 2220
ccggttctc caagcgggccc gggcggggga ggaggaagag ggctgggctg gagctagcaa 2280
ggggatattc ctctcccggc ttgagtcaga cgcggggcga tccgtcctcc cccgttccct 2340
cccaggagac gggaacttac ttcatttccc tggggcaggt tcgcccacgt taccaacttc 2400
tccccctccc ccagcaccac cgtcccttcc agcttccgcg cccccaccc aactgggcag 2460
gaccaggtc gtgctgccac cccctcttcg gggaaaggcg gccgcagccg cagacacctg 2520
ggggccgggg ctgggggtgg gggctcccta gcagccgccc gagcgttgtc caacacgtga 2580
gactcatgtg atgaagccgg gggagggcg gcaggtcgct ccttccctcc ccggcagtg 2640
ccagacgtgc ctggagtcac agggtagaac acgtagctcc aaccaccca ccgctccctc 2700
ccgctcccc gcccgccca cttctcattc acttggtcgc cagggcgag acagaccgag 2760
caggagcac acaccgccag tctgtgcgct gagtccgagc cagaggccgc ggggacaccg 2820

ggccatgcac gcccccaact gaagctgcat ctcaaagccg aagattccag cagcccaggg	2880
gatttcaaag agctcagact cagaggaaca tctgcgagaga gacccccgaa gccctctcca	2940
gggcagtcct catccagacg ctccgctagt gcagacagga gcgcgcagtg gccccggctc	3000
gccgcgccat ggagcggatc cccagcgcgc aaccaccccc cgcctgctg cccaaagcac	3060
cgggactgga gcacggagac ctaccaggga tgtaccctgc ccacatgtac caagtgtaca	3120
agtcaagacg gggaataaag cggagcgagg acagcaagga gacctacaaa ttgccgcacc	3180
ggctcatcga gaaaaagaga cgtgaccgga ttaacgagtg catcgcccag ctgaaggatc	3240
tcttaccgga acatctcaaa cttacaactt tgggtcactt ggaaaaagca gtggttcttg	3300
aacttacctt gaagcatgtg aaagcactaa caaacctaat tgatcagcag cagcagaaaa	3360
tcattgccct gcagagtggg ttacaagctg gtgagctgtc agggagaaat gtcgaaacag	3420
gtcaagagat gttctgtcga ggtttccaga catgtgcccg ggaggtgctt cagtatctgg	3480
ccaagcacga gaacactcgg gacctgaagt cttcgcagct tgtcacccac ctccaccggg	3540
tgggtctcga gctgctgcag ggtggtacct ccaggaagcc atcagaccca gctcccaaag	3600
tgatggactt caaggaaaaa cccagctctc cggccaaagg ttcggaaggc cctgggaaaa	3660
actgctgcc agtcatccag cggactttcg ctactcagag tggggagcag agcggcagcg	3720
acacggacac agacagtggc tatggaggag aatcggagaa gggcgacttg cgagtgagc	3780
agccgtgctt caaaagtgc cacggacgca ggttcacgat gggagaaagg atcggcgcaa	3840
ttaagcaaga gtccgaagaa cccccacaa aaaagaaccg gatgcagctt tcggatgatg	3900
aaggccattt cactagcagt gacctgatca gctccccgtt cctgggcccc caccacacc	3960
agcctcctt ctgcctgccc ttctacctga tcccaccttc agcgactgcc tacctgcccc	4020
tgctggagaa gtgctgggat cccacctcag tgccagtgt ataccaggc ctcaacgcct	4080
ctgccgcagc cctctctagc ttcatgaacc cagacaagat ctcggctccc ttgctcatgc	4140
cccagagact cccttctccc ttgccagctc atcgcgtccg cgactcttct gtcttgctcc	4200
aagctctgaa gccaatcccc cctttaaaact tagaaaccaa agactaaact ctctagggga	4260
tcctgtgctt ttgctttcct tcctcgctac ttcctaaaaa gcaacaaaaa agtttttgtg	4320
aatgctgcaa gattgttgca ttgtgtatac tgagataatc tgaggcatgg agagcagatt	4380
caggggtgtg gtgtgtgtgt gtgtgtgtgt gtgtatgtgc gtgtgctgac acatgtgtgc	4440
ctgcgtgttg gtataggact ttaaagctcc ttttggcata ggggaagtcac gaaggattgc	4500
ttgacatcag gagacttggg ggggattgta gcagacgtct gggcttttcc ccaccagag	4560
aatagcccc ttcgatacac atcagctgga ttttcaaaag cttcaaagtc ttggtctgtg	4620

49

```

agtcactctt cagtttggga gctgggtctg tggctttgat cagaaggtac tttcaaaaga 4680
gggctttcca gggctcagct cccaaccagc tgtaggacc ccaccctttt gcctttattg 4740
tcgacgtgac tcaccagacg tcggggagag agagcagtca gaccgagctt tctgctaaca 4800
tggggaggta gcaggcactg gcatagcacg gtagtggttt ggggagggtt ccgcaggctt 4860
gctccccacc cctgcctcgg aagaataaag agaatgtagt tccctactca ggctttcgta 4920
gtgattagct tactaaggaa ctgaaaatgg gcccttgta caagctgagc tgccccggag 4980
ggagggagga gttccctggg cttctggcac ctgtttctag gcctaaccat tagtacttac 5040
tgtgcaggga accaaaccaa ggtctgagaa atgcggacac cccgagcgag caccctaaag 5100
tgcacaaagc tgagtaaaaa gctgccccct tcaaacagaa ctagactcag ttttcaattc 5160
caccctaaaa ctctttttaa ccaagcttag cttctcaaag gcctaaccaa gccttggcac 5220
cgccagatcc tttctgtagg ctaattcttc ttgcccacg gcatatggag tgtccttatt 5280
gctaaaaagg attccgtctc cttcaaagaa gttttatttt tggccagag tacttgtttt 5340
cccgatgtgt ccagccagaa ccacaaatat cggggtacga gctc 5384

```

```

<210> 40
<211> 5356
<212> DNA
<213> Homo sapien

```

```

<400> 40
tgacatacga aaggttttta atatctacaa aagatgacag actgaaatag taacattctc 60
caggatggaa agatctatgc cacgtaggaa ttgtttgttc aaattattat tttttaaag 120
ccataagcta tggagagaat aacagagttt tgttttgttt tttaatcctg atacggtgga 180
agtctggtag gatgggaacc gttcttgaag cttgagaata cattcggaac aaataaaaag 240
taggaaatgc atacattttg ttattttacag ttgtagggca ggtaagggca aaacgtattg 300
ataaactcaa agaattttta gaatgatatt gtcatgagcc atgagccatg gaagctgcaa 360
ggtttgggct ccctcctgcc atcgggtggg ggggtgcggc ggaagggggg acgtgtcctg 420
tcctctccca gtgcttattt cctgctggaa ccgtgggttt gacgcagccg tgacagccct 480
tcaccttat taaaatcaca caatgaagca ggtcgcccaa gattgaaagc gcgtttccag 540
ggcggacccc ggtccccgcc tttggccagc aaaagtacat tctggcaggg agggatctct 600
gagccccgaa gtgaggtcat gccagttagc agttagaaaa ggaggcattt ctacacttg 660
ctaaagagtt tgacgcttgt gaaatataaa caggaagcgg gagctctcca aggcgacctg 720
aaggaaatac actgccagtt tggtaactt gtgaatgcca gttctctagg tcgtcaaggg 780
cactgccagt taagactgat cccctaactg caccggaggc tactactctt cgcccagctc 840

```

cggcagtgac cccatttttt tttcaatcct tcccacagcc aggtcactca gcagtggggt	900
gactgacccg aagcgcagtg ggaagggtt acagtcattct ctcaactccc cgaacttgca	960
gggcagggga gcgcagaggt cagtgaatc cccagccagg ccttatttta gtcattgagt	1020
tccactggct cctcctctg actgtatcag cgctgctcac accgcgggat tggagtttca	1080
gctgttttcc ctgctcccg ctgagtgaga aggtcacaga gcccggacct ttccggggga	1140
gggtgggtggg gccctccggc agcctaagt agattgcaga gatctggggc cgctggagga	1200
ggctccttcg ccccggtt ggggttttg gttagcttct ccccgcttc tctacctagg	1260
catctcggtt acaatgaaat taaaaacaaa acacaaagca acacaagttc tcattgtttt	1320
tccccctgt cagccagtgg cctcgataca attttcccag ggctgccttt cctttattgc	1380
tcaatttaac ctcttcctag aggttctgca agttcaaggg caggatccgt gtcggattca	1440
tcaactagt cagcgcctag tgcagagcag ggggtcaag aaatacgtgt cgaccgatg	1500
attgctaaga tttccctgta acaggcggtt ttttttctt gtctctctca tccctcctgc	1560
cctgggtgca gaccgcgaac agcgccgccc aacttcccaa caggcagga gagggccgat	1620
ccgggctggc acgccacgtc cccgggtct cagtccctgg gtcaacgcag agaggctcgc	1680
ggctccggt gctacgacct gagcccgcc acgtgaagcg tcacgccctg cgctgctgca	1740
gggcaggcgc cgccgcccgc gttgtgtaac gctgcagccg gaggggaggc gaccggggcg	1800
cctctgctcc ggagggggca gagggggag ggagcgccg gcacagcgcc cggtatcttt	1860
ctgcgcttga ctggccggga ggaaggggt gacctgggg caccagggg gctgcgcaac	1920
cggtgcccc cccgccgca ggttggtccc aggagcgag ctctgggag cgaggggccc	1980
tcggctcaca ccctccacgg tcagggtgcgc gcggcggtct cgggcgagc ctgagctgcc	2040
cctgcgcttg gccgtcgcc cgcttcccat ggggtgacat ccgccccgcc cctcggtccc	2100
tccccaggc gggcaattcc tggacgcgag ggtgagcagt gggggcaggg aagccaggac	2160
ggaaagaaac cccagcctct ggggaaggct ggggggcccg gcgactccct ccgcagcgcg	2220
ccggttcctc caagcgggcc gggcgggga ggaggaagag ggctgggctg gagctagcaa	2280
ggggatatcc ctctcccgcc ttgagtcaga cgcggcgga tccgtcctcc ccgttccct	2340
cccaggagac ggaacttac ttcatttccc tggggcaggt tcgcccacgt taccaacttc	2400
tccccctccc ccagcaccac cgctccctcc agcttccgcg cccccaccc aactgggcag	2460
gaccaggtc gtgctgccac cccctcttcg gggaaaggcg gccgcagccg cagacacctg	2520
ggggccgggg ctgggggtg gggctcccta gcagccgccc gagcgttgct caacacgtga	2580
gactcatgtg atgaagccgg gggagggcg gcaggtcgct ccttccctcc ccggcagtg	2640
ccagacgtgc ctggagtcac agggtagaac acgtagctcc aaccaccca ccgtccctc	2700

ccgcctcccc gcccgcccca cttctcattc acttggtctg cacggcgag acagaccgcg 2760
cagggagcac acaccgccag tctgtgcgct gagtcggagc cagaggccgc ggggacaccg 2820
ggccatgcac gcccccaact gaagctgcat ctcaaagccg aagattccag cagcccaggg 2880
gatttcaaa agctcagact cagaggaaca tctgcggaga gacccccgaa gccctctcca 2940
gggcagtcct catccagacg ctccgctagt gcagacagga gcgcgcagtg gccccggctc 3000
gccgcgccat ggagcggatc cccagcgcgc aaccaccccc cgcctgcctg cccaaagcac 3060
cgggactgga gcacggagac ctaccaggga tgtaccctgc ccacatgtac caagtgtaca 3120
agtcaagacg gggaataaag cggagcgagg acagcaagga gacctacaaa ttgccgcacc 3180
ggctcatcga gaaaaagaga cgtgaccgga ttaacgagtg catcgcccag ctgaaggatc 3240
tcttaccga acatctcaaa cttacaactt tgggtcactt ggaaaaagca gtggttcttg 3300
aacttacctt gaagcatgtg aaagcactaa caaaccta at tgatcagcag cagcagaaaa 3360
tcattgccct gcagagtggg ttacaagctg gtgagctgtc agggagaaat gtcgaaacag 3420
gtcaagagat gttctgctca ggtttccaga catgtgcccg ggaggtgctt cagtatctgg 3480
ccaagcacga gaacactcgg gacctgaagt cttcgcagct tgtcacccac ctccaccggg 3540
tggtctcgga gctgctgcag ggtggtacct ccaggaagcc atcagaccca gctcccaaag 3600
tgatggactt caaggaaaaa cccagctctc cggccaaagg ttcggaaggc cctgggaaaa 3660
actgcgtgcc agtcatccag cggactttcg ctactcagag tggggagcag agcggcagcg 3720
acacggacac agacagtggc tatggaggag aatcggagaa gggcgacttg cgcagtgagc 3780
agccgtgctt caaaagtgc cagcgacgca ggttcacgat gggagaaagg atcggcgcaa 3840
ttaagcaaga gtccgaagaa cccccacaa aaaagaaccg gatgcagctt tcggatgatg 3900
aaggccattt cactagcagt gacctgatca gctccccgtt cctgggcccc caccacacc 3960
agcctccttt ctgcctgccc ttctacctga tcccacctt agcgactgcc tacctgccc 4020
tgctggagaa gtgctggtat cccacctcag tgccagtgt ataccaggc ctcaacgcct 4080
ctgccgcagc cctctctagc ttcatgaacc cagacaagat ctcggtccc ttgctcatgc 4140
cccagagact cctttctccc ttgccagctc atccgtccgt cgactcttct gtcttctcc 4200
aagctctgaa gccaatcccc cttttaact tagaaaccaa agactaaact ctctagggga 4260
tcctgctgct ttgctttcct tcctcgctac ttctaaaaa gcaacaaaa agtttttg 4320
aatgctgcaa gattgttgca ttgtgtatac tgagataatc tgaggcatgg agagcagatt 4380
caggggtgtg gtgtgtgtgt gtgtgtgtgt gtgtatgtgc gtgtgcgtgc acatgtgtgc 4440
ctgcgtgttg gtataggact ttaaagctcc ttttggcata gggaagtcac gaaggattgc 4500

52

ttgacatcag gagacttggg ggggattgta gcagacgtct gggttttcc ccaccagag 4560
 aatagcccc ttcgatacac atcagctgga ttttcaaaag cttcaaagtc ttggctctgtg 4620
 agtcactctt cagtttggga gctgggtctg tggtttgat cagaaggtag tttcaaaaga 4680
 gggctttcca gggctcagct cccaaccagc tgtaggacc ccacctttt gcctttattg 4740
 tcgacgtgac tcaccagacg tcggggagag agagcagtc gaccgagctt tctgctaaca 4800
 tggggaggta gcaggcactg gcatagcacg gtagtggtt ggggagggtt ccgcaggctt 4860
 gctccccacc cctgcctcgg aagaataaag agaattgagt tccctactca ggctttcgt 4920
 gtgattagct tactaaggaa ctgaaaatgg gcccttgta caagctgagc tgccccggag 4980
 ggaggaggga gttccctggg cttctggcac ctgtttctag gcctaaccat tagtacttac 5040
 tgtgcaggga accaaaccaa ggtctgagaa atgcggacac cccgagcgag caccctaaag 5100
 tgcacaaagc tgagtaaaaa gctgccccct tcaaacagaa ctagactcag ttttcaattc 5160
 catcctaaaa ctcttttaa ccaagcttag cttctcaaag gcctaaccaa gccttggcac 5220
 cgccagatcc tttctgtagg ctaattcctc ttgccaacg gcatatggag tgccttatt 5280
 gctaaaaagg attcgtctc cttcaaagaa gttttattt tggtccagag tacctcgcc 5340
 gcgaccacgm taagcc 5356

<210> 41
 <211> 2176
 <212> DNA
 <213> Homo sapien

<400> 41
 cacgtgctcg agctgcttcc cgggcagctg gagcaggatg atagtgggccc aggcgtgacg 60
 agcgggcagt gtgctggggt aaaggacctg acaggcctca ggagggactt gcgcttcagg 120
 ccgggatcag gtgctgtgaa gctgcctgtg gagctggcct tggttttccg gaactcctcc 180
 agcttctgtc ggtagtactt gtacccttgg ctattgggct catacagaaa gctgaatgcc 240
 tggttctcac ggttgctctg gagggcaatg gtttccacct cgggaccccc gtccgctatg 300
 aacctggcca acttttctgc aagggtcttg acctcttctg cctctggggg tgaaacttta 360
 agcaggctgt cagtactggg ctcaactcac cacaccaca ccagcctcct ttctgcctgc 420
 ctttctacct gateccacct tcagcgactg cctacctgcc catgctggag aagtgtggt 480
 atcccacctc agtgccagtg ctataccag gcctcaacgc ctctgccga gccctctcta 540
 gcttcatgaa ccagacaag atctcggctc ccttgctcat gcccagaga ctcccttctc 600
 ccttgccagc tcatcgtcc gtcgactctt ctgtcttctg ccaagctctg aagccaatcc 660
 cccctttaa cttagaaacc aaagactaaa ctctctaggg gatcctgctg ctttgctttc 720

53

```

cttcctcgtc acttcctaaa aagcaacaaa aaagtttttg tgaatgctgc aagattgttg 780
catttgtgtat actgagataa tctgaggcat ggagagcaga ttcaggggtg gtgtgtgtgt 840
gtgtgtgtgt gtgtgtatgt gcgtgtgctg gcacatgtgt gcctgcgtgt tggatatagga 900
ctttaaagct ccttttgcca tagggaagtc acgaaggatt gcttgacatc aggagacttg 960
gggggggattg tagcagacgt ctgggctttt cccacccag agaatagccc ctttcgatac 1020
acatcagctg gatcttcaaa agcttcaaag tcttggctctg tgagtcactc ttcagtttgg 1080
gagctgggtc tgtggctttg atcagaaggt actttcaaaa gagggctttc cagggtctag 1140
ctcccaacca gctgttagga cccacccctt ttgcctttat tgtcgacgtg actcaccaga 1200
cgtcggggag agagagcagt cagaccgagc tttctgctaa catggggagg tagcaggcac 1260
tggcatagca cggtagtggt ttggggaggt ttccgcaggt ctgctcccca ccctgcctc 1320
ggaagaataa agagaatgta gttccctact caggctttcg tagtgattag cttactaagg 1380
aactgaaaat gggccccttg tacaagctga gctgccccgg agggagggag gagttccctg 1440
ggcttctggc acctgtttct aggcctaacc attagtactt actgtgcagg gaaccaaacc 1500
aaggtctgag aaatgcggac accccgagcg agcaccctaa agtgcacaaa gctgagtaaa 1560
aagctgcccc cttcaaacag aactagactc agttttcaat tccatcctaa aactcctttt 1620
aaccaagctt agcttctcaa aggcctaacc aagccttggc accgccagat cctttctgta 1680
ggctaattcc tcttgcccaa cggcatatgg agtgtcctta ttgctaaaaa ggattccgtc 1740
tccttcaaag aagttttatt tttgggtccag agtacttgtt ttcccgatgt gtccagccag 1800
ctccgcagca gcttttcaaa atgcactatg cctgattgct gatcgtgttt taactttttc 1860
ttttcctgtt tttatttttg tattaagtcg ttgcctttat ttgtaaagct gttataaata 1920
tatattatat aaatatatta aaaaggaaaa tgtttcagat gtttatttgt ataattactt 1980
gattcacaca gtgagaaaaa atgaatgtat tcctgttttt gaagagaaga ataatttttt 2040
ttttctctag ggagaggtag agtggtttata ttttgagacc ttctgaagg tgtaaaattg 2100
taaataatctt tatctatgag taaatgttaa gtagttgttt taaaatactt aataaaaaaa 2160
ttcttttctc gtggaa 2176

```

<210> 42
<211> 1047
<212> DNA
<213> Homo sapien

```

<400> 42
tgatctaagg ccaccctctc ggggagggag ttggggaagc tgggttggtg gggttggtag 60
ctcctaccta ctgtgtggca agaaggtatg ggtcatgaac agaaccaagg agctgcgctg 120

```

54

ctacagatgt taccacttct gtggctgcta cccactcct gggccgtccc tgaagtcct 180
 actccaatgt ggccagatga cctgcaaac cacacattcc tgcacacagt gtactgccag 240
 gatgggagtc ccagtgtggg actctctgag gcctacgacg aggaccagct tttcttcttc 300
 gacttttccc agaacactcg ggtgcctcgc ctgcccgaat ttgctgactg ggctcaggaa 360
 cagggagatg ctcttgccat tttatttgac aaagagtctt gcgagtggat gatccagcaa 420
 atagggccaa aacttgatgg gaaaatcccg gtgtccagag ggtttcctat cgctgaagtg 480
 ttcacgctga agccccgga gtttggcaag cccaacactt tggctctgtt tgctcagtaat 540
 ctcttccac ccatgctgac agtgaactgg cagcatcatt ccgtccctgt ggaaggattt 600
 gggcctactt ttgtctcagc tgcgatgga ctgagcttcc aggccttttc ttacttaaac 660
 ttcacaccag aaccttctga cattttctcc tgcattgtga ctacgaaat tgaccgctac 720
 acagcaattg cctattgggt accccggaac gcactgccct cagatctgct ggagaatgtg 780
 ctgtgtggcg tggcctttgg cctgggtgtg ctgggcatca tcgtgggcat tgttctcatc 840
 atctacttcc ggaagccttg ctgaggtgac tgattcttcc agaccagagt ttgatgccag 900
 cagcttcgga catccctggg gtgtgtgtgt agatttccca cctggggact ctgctgtccc 960
 tgggcttgca tcccagggat cccagagtgg cctgcctatc acaaccacat cccttcccc 1020
 cacaaggcaa taaatctcat ttcttta 1047

<210> 43
 <211> 871
 <212> DNA
 <213> Homo sapien

<400> 43
 ctctgatcgc cgatcacctc tgagaccacc ttgtcataa aaaaaatgcc catgttggtc 60
 ctctgccctg gacctgtgac attctggact atttctgtgt ttatttgtgg ccgagctcga 120
 gcgggccatt ttatttgaca aagagtcttg cgagtggatg atccagcaa tagggccaaa 180
 acttgatggg aaaatcccgg tgtccagagg gtttctatc gctgaagtgt tcacgtgaa 240
 gcccctggag tttggcaagc ccaacacttt ggtctgtttt gtcagtaate tcttcccacc 300
 catgctgaca gtgaactggc agcatcattc cgtccctgtg gaaggatttg ggcctacttt 360
 tgtctcagct gtcgatggac tcagcttcca ggcttttct tacttaaaact tcacaccaga 420
 accttctgac attttctcct gcattgtgac tcacgaaatt gaccgctaca cagcaattgc 480
 ctattgggta ccccggaacg cactgccctc agatctgctg gagaatgtgc tgtgtggcgt 540
 ggcccttggc ctgggtgtgc tgggcatcat cgtgggcatt gttctcatca tctacttccg 600
 gaagccttgc tcaggtgact gattcttcca gaccagagtt tgatgccagc agcttcggcc 660

55

atccaaacag aggatgctca gattttctcac atcctgccc a ggatctctc ttagggtaga 720
 agaagtctct gggacatccc tgggggtgtgt gtgtagattt cccacctggg gactctgctg 780
 tccctgggct tgcaccccag ggatcccaga gtggcctgcc tatcacaacc acatcccttc 840
 ccccccacaag gcaataaata tcatttcttt a 871

<210> 44

<211> 2754

<212> DNA

<213> Homo sapien

<400> 44

gcgggatggc aggggagcgc tccacgccct gtccccgtct aagctaccac ctttacttcc 60
 accaggctgg gaaccagggc ttccttgggt gtggctagca atgagtctgg atgacaattt 120
 gtcgggcacg agcggtatgg aagtggacga ccgctgtcg gcgctggagc agcggctgca 180
 gttacaggaa gacgagctgg cggctcctaaa ggcggcgctg gcggatgctc tgcgtcgcct 240
 gcgggcatgc gaagaacagg gagcggcgct acgcgcgcgg ggcacccccca agggccgggc 300
 gcctccgcgc ttaggcacca ctgcctcgggt gtgtcagctc ttgaaaggcc tccccaccag 360
 gacgccccctt aatggctcgg gacccccgcg gcgcgtgggt ggctatgcca cgtccccatc 420
 ctctcccaag aaggaggcga cctccgggcg cagcagtgtc cgcgctact tgtcaccaga 480
 gcgcctcgcc tcgggtgcgc gtgaggacct ccgcagccgg accacatcct ccagcagcaa 540
 ctgtagcgcc aaaaagggaag gcaaaaccaa agaagttatc ttcagtgtgg aggatggctc 600
 cgtgaaaatg ttcctgaggg gccgcctgt gccatgatg atcccagacg agctggcacc 660
 cacctacagc ctggacacac gctcggagct gccttcttgc cggctcaagc tggagtgggt 720
 ctatggctac cgtggccgag actgccgggc caaccttat ttgtgccc cgggggagat 780
 agtgtacttt gtggcctccg tagccgtgct atacagcgtg gaggagcaga ggcagcgaca 840
 ctacctggga cacaacgatg acatcaaata cttggccatc caccagata tggtcacat 900
 cgccacggga cagggtggcg gaaccactaa ggaagggaag ccgctgccgc cccacgtgcg 960
 catctgggac tcagtttccc tctccacctt acacgtgctg ggcttggggg tgtttgacag 1020
 agcctgtgtc tgtgtgggct tctccaaatc taatggaggc aacctgctgt gtgcagtggg 1080
 tgaatccaat gatcacatgc tctcgggtgt ggactgggcc aaggagacca aggtgggtga 1140
 tgtcaagtgc tccaatgagg ctgtattggt ggccaccttc caccacacgg accccactgt 1200
 gcttatcacc tgcgggaaat ctacatcta cttctggacc ttggaggggg gcagcttgag 1260
 caagcggcaa ggcctctttg agaaacatga gaaaccgaag tatgtgctgt gtgtgacctt 1320
 tttggaaggt ggcgacgtgg tcacggggga ctctgggggg aacctctatg tttggggcaa 1380

56

aggtgggaac cgtatcacac aggcgggtgct gggcgccac gacggcggcg tgtttgggct 1440
 ctgcgccctg cgggacggga cgctgggtgtc tggagggggc cgtgatcggc gggtggtcct 1500
 ctgggggttct gactacagca agctgcagga agtggagggtc cctgaggact ttggccctgt 1560
 gcgcaccgtg gcagagggcc acggagacac actgtacgtg gggaccaccc gcaattccat 1620
 cctgcagggc tccgtgcaca caggcttctc actgctggtc cagggccatg tggaagagct 1680
 gtggggcctg gccacacacc ccagtcgggc ccagtttgtg acctgcgggc aggataagct 1740
 ggtgcattcta tggagctcag attcccacca gcccctgtgg agcaggatca tcgaggaccc 1800
 tgcccgtca gccggcttcc acccagtggt ctctgtcctg gctgtgggta cagtgactgg 1860
 cagatggctg ctgctggaca cggagacca tgacctgggt gctatccaca cagacggcaa 1920
 tgaacagatc tcagtgttca gcttctcccc agacggggcg tacctggccg tgggtccca 1980
 cgacaacttg gtgtacgtgt acacggtgga ccagggcggc cgcaagggtca gccgcctggg 2040
 caagtgtctg ggccattcca gttttatcac ccacctggat tggggccagg acagcagctg 2100
 ctttgtcacc aactccgggg actatgagat tctgtactgg gacccggcta cctgtaagca 2160
 gatcaccagt gcggatgctg tgaggaacat ggaatgggcc acagctactt gtgtcctagg 2220
 gtttgggggtg tttgggatct ggtctgaggg ggcggacggc actgatatac acgctgtggc 2280
 ccgctctcat gatgggaagt tgctggcttc agctgatgac tttggcaaag ttacacctgt 2340
 tagctacccc tgetgtcagc ctcgagccct cagccacaag tacggtggac acagcagcca 2400
 tgtgacaaat gtggccttct tgtgggatga cagcatggcc ctgaccacag ggggcaagga 2460
 caccagtgtg ctacagtggc ggggtggtctg atgcggccag ggaagagtca ggtgtcaggg 2520
 caggaattct attttcggga gatgtctatt gccgagtaga gtaatatata ccagagtat 2580
 gtctatagca gaggggggta tgggggcggg agggtagact gacatacaga agtctctatt 2640
 tatccgggtg ggaagagggg gtcacatcgc tttggggatc cattgggtgt tggtttgggg 2700
 tgtttttttaa gttttttctt ttatatcatc cagaaataaa gacacgtaca ctaa 2754

<210> 45
 <211> 2309
 <212> DNA
 <213> Homo sapien

<400> 45
 ctcggccccc ttttcatcgg agacctccgg ggagcgggtg ggggtggagga atgggtttctc 60
 cccttttctg aactgaatac taagaccctt tttttttctt tgccttttcc tgacagcaaa 120
 accaaagaag ttatcttcag tgtggaggat ggctccgtga aaatgttctt gaggggcccgc 180
 cctgtgcccc tgatgatccc agacgagctg gcacccacct acagcctgga cacacgctcg 240

57

gagctgcctt cttgccggct caagctggag tgggtctatg gctaccgtgg ccgagactgc	300
cgggccaaacc tttatttgcg gccacccggg gagatagtgt actttgtggc ctccgtagcc	360
gtgctataca gcgtggagga gcagaggcag cgacactacc tgggacacaa cgatgacatc	420
aaatgcttgg ccatccaccc agatatggtc accatcgcca cgggacaggt ggcgggaacc	480
actaaggaag ggaagccgct gccgccccac gtgcgcacatc gggactcagt ttccctctcc	540
accttacacg tgctgggctt gggggtgttt gacagagccg tgtgctgtgt gggcttctcc	600
aaatctaata gaggcaacct gctgtgtgca gtggatgaat ccaatgatca catgctctcg	660
gtgtgggact gggccaagga gaccaagggt gtggatgtca agtgctcaa tgaggctgta	720
ttggtggcca ccttccaccc cacggacccc actgtgctta tcacctgcgg gaaatctcac	780
atctacttct ggaccttga ggggggcagc ttgagcaagc ggcaaggcct ctttgagaaa	840
catgagaaac cgaagtatgt gctgtgtgtg accttttttg aagggtggcg cgtggtcacg	900
ggggactctg gggggaacct ctatgttttg ggcaaagggt ggaaccgtat cacacaggcg	960
gtgctgggcg cccacgacgg cggcgtgttt gggctctgcg ccctgcggga cgggacgctg	1020
gtgtctggag ggggccgtga tcggcgggtg gtcctctggg gttctgacta cagcaagctg	1080
caggaagtgg aggtccctga ggactttggc cctgtgcgca ccgtggcaga gggccacgga	1140
gacacactgt acgtggggac caccgcgaat tccatcctgc agggctccgt gcacacaggc	1200
ttctcactgc tgggccaggg ccatgtggaa gagctgtggg gcctggccac acacccagct	1260
cgggcccagt ttgtgacctg cgggcaggat aagctgggtg atctatggag ctgagattcc	1320
caccagcccc tgtggagcag gatcatcgag gaccctgccc gctcagccgg cttccacccc	1380
agtggctctg tcctggctgt gggtagctg actggcagat ggctgctgct ggacacggag	1440
acctatgacc tgggtggctat ccacacagac ggcaatgaac agatctcagt ggtcagcttc	1500
tccccagacg gggcgctacct ggccgtgggc tcccacgaca acttgggtga cgtgtacacg	1560
gtggaccagg gcggccgcaa ggtcagccgc ctgggcaagt gtcggggcca ttccagtttt	1620
atcacccacc tggattgggc ccaggacagc agctgctttg tcaccaactc cggggactat	1680
gagattctgt actgggaccc ggctacctgt aagcagatca ccagtgcgga tgctgtgagg	1740
aacatggaat gggccacagc tacttgtgtc ctagggtttg ggggtgtttg gatctggtct	1800
gagggggcg acggcactga tatcaacgct gtggcccgct ctcatgatgg gaagtgtgtg	1860
gcttcagctg atgacttttg caaagttcac ctgttttagct acccctgctg tcagcctcga	1920
gccctcagcc acaagtacgg tggacacagc agccatgtga caaatgtggc cttcttgtgg	1980
gatgacagca tggccctgac cacagggggc aaggacacca gtgtgctaca gtggcgggtg	2040
gtctgatgcg gccagggaag agtcagggtg cagggcagga attctatttt cgggagatgt	2100

ctattgccga gtagagtaat atataccag agtatgtcta tagcagaggg ggttatgggg 2160
gcgggagggg agactgacat acagaagtct ctatttatcc ggggtgggaag agggagtcac 2220
atcgcttttg ggatccattg gtgtttgggt tgggggtgtt ttttaagttt ttcttttata 2280
tcatccagaa ataaagacac gtacactaa 2309

<210> 46
<211> 2312
<212> DNA
<213> Homo sapien

<400> 46
ctccaccccc gggccgaag ctccgccacc cgccgccatg agtagctttg gagctgggtg 60
agaccctttc gcagaccacc ccccatctcc tctccgcgct cccggggctt agatctcagc 120
aaaaccaaag aagttatctt cagtgtggag gatggctccg tgaaaatgtt cctgaggggc 180
cgccctgtgc ccatgatgat ccagacgag ctggcaccca cctacagcct ggacacacgc 240
tcggagctgc cttcttgccg gctcaagctg gagtgggtct atggctaccg tggccgagac 300
tgccggggcca acctttatctt gctgccacc ggggagatag tgtactttgt ggcctccgta 360
gccgtgctat acagcgtgga ggagcagagg cagcgacact acctgggaca caacgatgac 420
atcaaagtct tggccatcca ccagatatg gtcaccatcg ccacgggaca ggtggcggga 480
accactaagg aagggaagcc gctgccgccc cactgcgca tctgggactc agtttccctc 540
tccaccttac acgtgctggg cttgggggtg tttgacagag ccgtgtgctg tgtgggcttc 600
tccaaatcta atggaggcaa cctgctgtgt gcagtggatg aatccaatga tcacatgctc 660
tcgggtgtggg actgggcca ggagaccaag gtggtggatg tcaagtgtc caatgaggct 720
gtattggtgg ccaccttcca cccacggac cccactgtgc ttatcacctg cgggaaatct 780
cacatctact tctggacctt ggaggggggc agcttgagca agcggcaagg cctctttgag 840
aaacatgaga aaccgaagta tgtgtgtgt gtgacctttt tgggaagggtg cgacgtggtc 900
acgggggact ctggggggaa cctctatgtt tggggcaaag gtgggaaccg tatcacacag 960
gcggtgctgg gcgcccacga cggcggcgtg tttgggtctc gcgcctgctg ggacgggacg 1020
ctggtgtctg gagggggccg tgatcggcgg gtggtcctct ggggttctga ctacagcaag 1080
ctgcaggaag tggagggtcc tgaggacttt ggccctgtgc gcaccgtggc agagggccac 1140
ggagacacac tgtacgtggg gaccaccgc aattccatcc tgacgggtc cgtgcacaca 1200
ggcttctcac tgctgggtcca gggccatgtg gaagagctgt ggggcctggc cacacacccc 1260
agtcggggcc agtttgtgac ctgcgggcag gataagctgg tgcattatg gagctcagat 1320
tcccaccagc ccctgtggag caggatcctc gaggaccctg cccgctcagc cggcttccac 1380


```

cccagtggtgct ctgtcctggc tgtgggtaca gtgactggca gatggctgct gctggacacg 1440
gagaccccatg acctggtggc tatccacaca gacggcaatg aacagatctc agtggtcagc 1500
ttctccccag acggggcgta cctggccgtg ggctcccacg acaacttggg gtacgtgtac 1560
acggtggacc agggcgggcg caaggtcagc cgcctgggca agtgctcggg ccattccagt 1620
tttatcaccc acctggattg ggcccaggac agcagctgct ttgtcaccaa ctccggggac 1680
tatgagattc tgtactggga cccggctacc tgtaagcaga tcaccagtgc ggatgctgtg 1740
aggaacatgg aatggggcac agctacttgt gtcctagggg ttgggggtgtt tgggatctgg 1800
tctgaggggg cggaacggcac tgatatcaac gctgtggccc gctctcatga tgggaagtgt 1860
ctggcttcag ctgatgactt tggcaaagtt cacctgttta gctaccctg ctgtcagcct 1920
cgagccctca gccacaagta cgggtggacac agcagccatg tgacaaatgt ggccttcttg 1980
tgggatgaca gcatggccct gaccacaggg ggcaaggaca ccagtgtgct acagtggcgg 2040
gtggtctgat gcggccaggg aagagtcagg tgtcagggca ggaattctat ttctgggaga 2100
tgtctattgc cgagtagagt aatatatacc cagagtatgt ctatagcaga gggggttatg 2160
ggggcgggag ggtagactga catacagaag tctctattta tccgggtggg aagagggagt 2220
cacatcgctt tggggatcca ttggtgtttg gtttgggggtg ttttttaagt tttttctttt 2280
atatcatcca gaaataaaga cacgtacact aa 2312

```

```

<210> 47
<211> 2826
<212> DNA
<213> Homo sapien

```

```

<400> 47
ggccagacgg tacctagggg gatccggggc cagcaatcct ggccctagga ttgtgttggg 60
gggcttgga ggcggtgtcc cgacccccct ttctgggcat cctggggatg accccagcgc 120
cgggcccggc gccggccgcc tgatgccagg cgagctgagc tggggatgct ggaacgaagg 180
gcgttgctat ggcaacggga ggcagggcct ggttgggggg accgggcccg ggctgggacc 240
ggggggggcg gcggtggctg tgggtggggc atggcggagc gcggcccggc cttctcgggc 300
ctgtacgaca cgtcctcgct gctgcgatac tgcaacgatg acaatttgtc gggcacgagc 360
ggatggaag tggacgaccg cgtgtcggcg ctggagcagc ggctgcagtt acaggaagac 420
gagctggcgg tcctaaaggc ggcgctggcg gatgctctgc gtcgcctgcg ggcagcgaa 480
gaacagggag cggcgctacg cgcgcggggc acccccaagg gccgggcgcc tccgcgctta 540
ggcaccactg cctcggtgtg tcagctcttg aaaggccttc ccaccaggac gcccttaat 600
ggctcgggac ccccgcgggc cgtgggtggc tatgccacgt ccccatcctc tcccaagaag 660

```

gaggcgacct ccgggcgcag cagtgtccgc cgctacttgt caccagagcg cctcgacctg	720
gtgcgcctg aggacccccg cagccggacc acatcctcca gcagcaactg tagcgccaaa	780
aaggaaaggca aaaccaaaga agttatcttc agtgtggagg atggctccgt gaaaatgttc	840
ctgagggggcc gccctgtgcc catgatgac ccagacgagc tggcacccac ctacagcctg	900
gacacacgct cggagctgcc ttcttgccgg ctcaagctgg agtgggtcta tggctaccgt	960
ggccgagact gccgggcca cctttatttg ctgccaccg gggagatagt gtactttgtg	1020
gcctccgtag ccgtgctata cagcgtggag gagcagaggc agcgacacta cctgggacac	1080
aacgatgaca tcaaagtctt ggccatccac ccagatatgg tcaccatcgc cacgggacag	1140
gtggcgggaa ccactaagga agggaagccg ctgccgccc acgtgcgcat ctgggactca	1200
gtttccctct ccaccttaca cgtgctgggc ttgggggtgt ttgacagagc cgtgtgctgt	1260
gtgggcttct ccaaactctaa tggaggcaac ctgctgtgtg cagtggatga atccaatgat	1320
cacatgctct cgggtgtgga ctgggccaag gagaccaagg tgggtgatgt caagtgtcc	1380
aatgaggctg tattggtggc caccttcac ccacggacc ccactgtgt tatcacctgc	1440
gggaaatctc acatctactt ctggaccttg gaggggggca gcttgagcaa gcggcaaggc	1500
ctctttgaga aacatgagaa accgaagtat gtgctgtgtg tgacctttt ggaagggtggc	1560
gacgtgttca cgggggactc tggggggaac ctctatgttt ggggcaaagg tccctgagga	1620
ctttggccct gtgcgcaccg tggcagaggg ccacggagac aactgtacg tggggaccac	1680
ccgcaattcc atcctgcagg gctccgtgca cacaggcttc tactgtctgg tccagggcca	1740
tgtggaagag ctgtggggcc tggccacaca cccagtcg gcccagtttg tgacctgcg	1800
gcaggataag ctggtgcac tatggagctc agattccac cagcccctgt ggagcaggat	1860
catcgaggac cctgcccgt cagccggctt ccacccagt ggctctgtcc tggctgtggg	1920
tacagtgact ggcagatggc tgctgtgga cagggagacc catgacctgg tggctatcca	1980
cacagacggc aatgaacaga tctcagtgg cagcttctcc ccagacggg cgtacctggc	2040
cgtgggctcc cagacaact tgggtgtacgt gtacacgggt gaccagggcg gccgcaagg	2100
cagccgcctg ggcaagtgt cgggccattc cagttttatc acccacttg attgggcca	2160
ggacagcagc tgctttgtca ccaactccg ggactatgag attctgtact gggacccggc	2220
tacctgtaag cagatcacca gtgcggatgc tgtgaggaac atggaatgg ccacagctac	2280
ttgtgtccta gggtttggg tgtttgggat ctggtctgag gggcgggacg gactgatata	2340
caacgctgtg gcccgctctc atgatgggaa gtgtctggct tcagctgatg actttggcaa	2400
agttcacctg tttagctacc cctgctgtca gcctcgagcc ctcagccaca agtacgggtg	2460

61

acacagcagc catgtgacaa atgtggcctt cttgtgggat gacagcatgg ccctgaccac 2520
 aggggggcaag gacaccagtg tgctacagtg gcggtgggtc tgatgcggcc agggaagagt 2580
 caggtgtcag ggcaggaatt ctatcttcgg gagatgtcta ttgccgagta gagtaatatata 2640
 taccagagt atgtctatag cagagggggt tatgggggcg ggagggtaga ctgacataca 2700
 gaagtctcta tttatccggg tgggaagagg gaggcacatc gctttgggga tccattgggtg 2760
 tttggtttgg ggtgtttttt aagttttttc ttttatatca tccagaaata aagacacgta 2820
 cactaa 2826

<210> 48
 <211> 2118
 <212> DNA
 <213> Homo sapien

<400> 48
 ctccaccccc ggccccgaag ctccgccacc cgccgccatg agtagctttg gagctggcaa 60
 aaccaaagaa gttatcttca gtgtggagga tggctccgtg aaaatgttcc tgaggggccc 120
 ccctgtgccc atgatgatcc cagacgagct ggcacccacc tacagcctgg acacacgctc 180
 ggagctgcct tcttgccggc tcaagctgga gtgggtctat ggctaccgtg gccgagactg 240
 ccgggccaac ctttatttgc tgcccaccgg ggagatagtg tactttgtgg cctccgtagc 300
 cgtgctatac agcgtggagg agcagaggca gcgacactac ctgggacaca acgatgacat 360
 caaatgcttg gccatccacc cagatatggt caccatcgcc acgggacagg tggcgggaac 420
 cactaaggaa ggaagccgc tgccgccccca cgtgcgcac cgtggactcag tttccctctc 480
 caccttacac gtgctgggct tgggggtgtt tgacagagcc gtgtgctgtg tgggcttctc 540
 caaatctaata ggaggcaacc tgctgtgtgc agtggatgaa tccaatgatc acatgctctc 600
 ggtgtgggac tgggccaagg agaccaagg ggtggatgtc aagtgtcca atgaggctgt 660
 attggtggcc accttccacc ccacggaccc cactgtgctt atcacctgcg ggaaatctca 720
 catctacttc tggaccttgg aggggggag cttgagcaag cggcaaggcc tctttgagaa 780
 acatgagaaa ccgaagtatg tgctgtgtgt gacctttttg gaagggtggcg acgtggtcac 840
 gggggactct ggggggaacc tctatgtttg gggcaaagg gggaaaccgta tcacacaggc 900
 ggtgctgggc gccacgacg gcggcgtgtt tgggctctgc gccctgcggg acgggacgct 960
 ggtgtctgga gggggccgtg atcggcgggt ggtcctctgg ggttctgact acagcaagct 1020
 gcaggaagtg gaggtccctg aggactttgg ccctgtgcgc accgtggcag agggccacgg 1080
 agacacactg tacgtgggga cccccgaa ttccatcctg cagggtccg tgcacacagg 1140
 cttctcactg ctggtccagg acctgtcccg ctgagccggc ttccaccca gtggctctgt 1200

62

cctggctgtg ggtacagtga ctggcagatg gctgctgctg gacacggaga cccatgacct 1260
 ggtggctatc cacacagacg gcaatgaaca gatctcagtg gtcagcttct cccacagacgg 1320
 ggcgtacctg gccgtgggct cccacgacaa cttggtgtac gtgtacacgg tggaccaggg 1380
 cggccgcaag gtcagccgcc tgggcaagtg ctctgggcat tccagtttta tcaccacct 1440
 ggattgggccc caggacagca gctgctttgt caccaactcc ggggactatg agattctgta 1500
 ctgggacccg gctacctgta agcagatcac cagtgcggat gctgtgagga acatggaatg 1560
 ggccacagct acttgtgtcc tagggtttgg ggtgtttggg atctggtctg agggggcgga 1620
 cggcactgat atcaacgctg tggcccgcct tcctgatggg aagtgtctg ctccagctga 1680
 tgactttggc aaagtccacc tgtttagcta cccctgctgt cagcctcgag ccctcagcca 1740
 caagtacggt ggacacagca gccatgtgac aaatgtggcc ttcttgtggg atgacagcat 1800
 ggccctgacc acagggggca aggacaccag tgtgctacag tggcgggtgg tctgatgcgg 1860
 ccaggggaaga gtcaggtgtc agggcaggaa ttctattttc gggagatgtc tattgccgag 1920
 tagagtaata tataccaga gtatgtctat agcagagggg gttatggggg cgggagggta 1980
 gactgacata cagaagtctc tatttatccg ggtgggaaga gggagtcaca tcgctttggg 2040
 gatccattgg tgtttggttt ggggtgtttt ttaagttttt tcttttatat catccagaaa 2100
 taaagacacg tacactaa 2118

<210> 49
 <211> 2152
 <212> DNA
 <213> Homo sapien

<400> 49
 ctccaccccc ggccccgaag ctccgccacc cgccgccatg agtagctttg gagctggcaa 60
 aaccaaagaa gttatcttca gtgtggagga tggctccgtg aaaatgttcc tgaggggccc 120
 ccctgtgccc atgatgatcc cagacgagct ggcacccacc tacagcctgg acacacgctc 180
 ggagctgcct tcttgccggc tcaagctgga gtgggtctat ggctaccgtg gccgagactg 240
 ccgggccaac ctttatttgc tgcccaccgg ggagatagtg tactttgtgg cctccgtagc 300
 cgtgctatac agcgtggagg agcagaggca gcgacactac ctgggacaca acgatgacat 360
 caaatgcttg gccatccacc cagatatggt caccatcgcc acgggacagg tggcgggaac 420
 cactaaggaa gggaagccgc tgccgcccc cgtgcgcac tgggactcag tttccctctc 480
 caccttacac gtgctgggct tgggggtgtt tgacagagcc gtgtgctgtg tgggcttctc 540
 caaatctaata ggaggcaacc tgctgtgtgc agtggatgaa tccaatgatc acatgctctc 600
 ggtgtgggac tgggccaagg agaccaaggt ggtggatgtc aagtgtctca atgaggctgt 660

63

attggtggcc accttccacc ccacggaccc cactgtgctt atcacctgcg ggaaatctca 720
 catctacttc tggaccttgg aggggggag cttgagcaag cggcaaggcc tctttgagaa 780
 acatgagaaa ccgaagtatg tgctgtgtgt gacctttttg gaaggtggcg acgtggtcac 840
 gggggactct ggggggaacc tctatgtttg gggcaaagggt gggaaaccgta tcacacaggc 900
 ggtgctgggc gccacgacg gcggcggtt tgggctctgc gccctgcggg acgggacgct 960
 ggtgtctgga gggggccgtg atcggcggtt ggtcctctgg ggttctgact acagcaagct 1020
 gcaggaagtg gaggtccctg aggacttttg ccctgtgcgc accgtggcag agggccacgg 1080
 agacacactg tacgtgggga ccaccgcaa ttccatcctg cagggtccg tgcacacagg 1140
 cttctcactg ctggtccagg gccatgtgga agagctgtgg ggcctggcca cacacccag 1200
 tcggggccag tttgtgacct gcgggcagga taagctggtg catctatgga gctcagattc 1260
 ccaccagccc ctgtggagca ggatcatcga ggacctgcc cgctcagccg gcttccaccc 1320
 cagtggctct gtcttggtg tgggtacagt gactggcaga tggctgctgc tggacacgga 1380
 gacctatgac ctggtggcta tccacacaga cggcaatgaa cagatctcag tggtcagctt 1440
 ctccccaggg ccattccagt tttatcacc acctggattg ggcccaggac agcagctgct 1500
 ttgtcaccaa ctccggggac tatgagattc tgtactggga cccggctacc tgtaagcaga 1560
 tcaccagtgc ggatgctgtg aggaacatgg aatggggcac agctacttgt gtcctaggg 1620
 ttgggggtgt tgggatctg tctgagggg cgagcggcac tgatatcaac gctgtggccc 1680
 gctctcatga tgggaagtgt ctggcttcag ctgatgactt tggcaaagtt cacctgttta 1740
 gctacccctg ctgtcagcct cgagccctca gccacaagta cgggtggacac agcagccatg 1800
 tgacaaatgt ggccttcttg tgggatgaca gcatggccct gaccacaggg ggcaaggaca 1860
 ccagtgtgct acagtggcgg gtggtctgat gcggccaggg aagagtcagg tgtcagggca 1920
 ggaattctat ttccgggaga tgtctattgc cgagtagagt aatatatacc cagagtatgt 1980
 ctatagcaga gggggttatg ggggcgggag ggtagactga catacagaag tctctattta 2040
 tccgggtggg aagagggagt cacatcgctt tggggatcca ttggtgtttg gtttggggtg 2100
 ttttttaagt tttttctttt atatcatcca gaaataaaga cacgtacact aa 2152

<210> 50
 <211> 1899
 <212> DNA
 <213> Homo sapien

<400> 50
 aggctaggta gcgggctggg tgtgaatcgg ccgaggtcgc ccgggacagg tggcgggaac 60
 cactaaggaa gggaagccgc tgccgcccc cgtgcgcac tgggactcag tttccctctc 120

64

caccttacac	gtgctgggct	tgggggtgtt	tgacagagcc	gtgtgctgtg	tgggcttctc	180
caaatcttgc	tccaatgagg	ctgtattggt	ggccaccttc	cacccacagg	acccactgt	240
gcttatcacc	tgcgggaaat	ctcacatcta	cttctggacc	ttggaggggg	gcagcttgag	300
caagcggcaa	ggcctctttg	agaaacatga	gaaaccgaag	tatgtgctgt	gtgtgacctt	360
tttgaaggt	ggcgacgtgg	tcacggggga	ctctgggggg	aacctctatg	tttggggcaa	420
agggtgggaac	cgtatcacac	aggcggtgct	gggcgcccac	gacggcggcg	tgtttgggct	480
ctgcgccctg	cgggacggga	cgctgggtgc	tggagggggc	cgtgatcggc	gggtggctct	540
ctgggggttct	gactacagca	agctgcagga	agtggaggtc	cctgaggact	ttggccctgt	600
gcgcaccgtg	gcagagggcc	acggagacac	actgtacgtg	gggaccaccc	gcaattccat	660
cctgcagggc	tccgtgcaca	caggcttctc	actgctggtc	caggatccag	caaccaagag	720
tttaactcca	agcacagcag	aggggcccga	ggcccctgct	ccaactgtgc	tgccctctgc	780
cactctgatt	gggggtggca	cccttcaggg	ccatgtggaa	gagctgtggg	gcctggccac	840
acaccccagt	cgggcccagt	ttgtgacctg	cgggcaggat	aagctggtgc	atctatggag	900
ctcagattcc	caccagcccc	tgtggagcag	gatcatcgag	gacctgccc	gctcagccgg	960
cttccacccc	agtggctctg	tcctggctgt	gggtacagtg	actggcagat	ggctgctgct	1020
ggacacggag	acccatgacc	tggtggctat	ccacacagac	ggcaatgaac	agatctcagt	1080
ggtcagcttc	tcccagacg	gggcgtacct	ggcgtgggc	tcccacgaca	acttgggtga	1140
cgtgtacacg	gtggaccagg	gcggccgcaa	ggtcagccgc	ctgggcaagt	gctcgggcca	1200
ttccagtttt	atcacccacc	tggattgggc	ccaggacagc	agctgctttg	tcaccaactc	1260
cggggactat	gagattctgt	actgggaccc	ggctacctgt	aagcagatca	ccagtgcgga	1320
tgctgtgagg	aacatggaat	gggccacagc	tacttgtgtc	ctagggtttg	gggtgtttgg	1380
gatctggtct	gagggggcgg	acggcactga	tatcaacgct	gtggcccgc	ctcatgatgg	1440
gaagttgctg	gcttcagctg	atgactttgg	caaagttcac	ctgttttagct	acccctgctg	1500
tcagcctcga	gccctcagcc	acaagtacgg	tggacacagc	agccatgtga	caaatgtggc	1560
cttcttgtgg	gatgacagca	tggccctgac	cacagggggc	aaggacacca	gtgtgctaca	1620
gtggcggggt	gtctgatgcg	gccagggaag	agtcagggtg	cagggcagga	attctatttt	1680
cgggagatgt	ctattgccga	gtagagtaat	atatacccag	agtatgtcta	tagcagaggg	1740
ggttatgggg	gcgggagggg	agactgacat	acagaagtct	ctatttatcc	gggtgggaag	1800
agggagtccac	atcgcttttg	ggatccattg	gtgtttgggt	tgggggtgtt	tttaagtttt	1860
ttctttttata	tcatccagaa	ataaagacac	gtacactaa			1899

65

<210> 51
 <211> 2951
 <212> DNA
 <213> Homo sapien

<400> 51
 ttcctrcggs cattgcgsag atttcatctg gccaggacac tggctgtcca cctggcactg 60
 gtcccgacag aggccgagct ggggaaagt ratgttact gggggcagga accctcctta 120
 tctccaggct gcttcccaac ctccaggagg cacaagatta gcagaaagcc acgttgagtc 180
 agccagtaac atggagcaac tgacaaggga aactgaggac tattccaaac aagccctctc 240
 actggtgctc aaggccctgc atgaaggagt cggaagcgga agcggtagcc cggacggtgc 300
 tgtggtgcaa gggcttgtgg aaaaattgga gaaaaccaag tccctggccc agcagttgac 360
 aagggaggcc actcaagcgg aaattgaagc agataggtct tatcagcaca gtctccgcct 420
 cctggattca gtgtctcggc ttcaggaggt cagtgatcag tcctttcagg tggaagaagc 480
 aaaggagatc aaacaaaaag cggattcact ctcaagcctg gtaaccaggc atatggatga 540
 gttcaagcgt acacagaaga atctgggaaa ctggaaagaa gaagcacagc agctcttaca 600
 gaatggaaaa agtgggagag agaaatcaga tcagctgctt tcccgtgcca atcttgctaa 660
 aagcagagca caagaagcac tgagtatggg caatgccact ttttatgaag ttgagagcat 720
 ccttaaaaac ctccagaggt ttgacctgca ggtggacaac agaaaagcag aagctgaaga 780
 agccatgaag agactctcct acatcagcca gaaggtttca gatgccagtg acaagacca 840
 gcaagcagaa agagccctgg ggagcgctgc tgctgatgca cagagggcaa agaatggggc 900
 cggggaggcc ctggaaatct ccagtgatg tgaacaggag attgggagtc tgaacttgga 960
 agccaatgtg acagcagatg gagccttggc catggaaaag ggactggcct ctctgaagag 1020
 tgagatgagg gaagtggaag gagagctgga aagggaaggag ctggagtttg acacgaatat 1080
 ggatgcagta cagatggtga ttacagaagc ccagaagggt gataccagag ccaagaacgc 1140
 tggggttaca atccaagaca cactcaacac attagacggc ctctgcac tcatggacca 1200
 gcctctcagt gtagatgaag aggggctggt ctactggag cagaagcttt cccgagccaa 1260
 gaccagatc aacagccaac tgcggcccat gatgtcagag ctggaagaga gggcacgtca 1320
 gcagaggggc cacctccatt tgctggagac aagcatagat gggattctgg ctgatgtgaa 1380
 gaacttgag aacattaggg acaacctgcc cccaggctgc tacaataccc aggtcttga 1440
 gcaacagtga agctgccata aatatttctc aactgaggtt cttgggatac agatctcagg 1500
 gctcgggagc catgtcatgt gagtgggtgg gatggggaca tttgaacatg tttaatgggt 1560
 atgctcaggt caactgacct gacccattc ctgatcccat ggccagggtg ttgtcttatt 1620
 gcaccatact ccttgcttcc tgatgctggg caatgaggca gatagcactg ggtgtgagaa 1680

tgatcaagga tctggacccc aaagaataga ctggatggaa agacaaactg cacaggcaga 1740
 tgtttgcctc ataatagtcg taagtggagt cctggaattt ggacaagtgc tgttgggata 1800
 tagtcaactt attctttgag taatgtgact aaaggaaaaa actttgactt tgcccaggca 1860
 tgaaattctt cctaattgtca gaacagagtgc caaccagtc aactgtggc cagtaaaata 1920
 ctattgcctc atattgtcct ctgcaagctt cttgctgac agagtctctc ctacttacia 1980
 cccagggtgt gaacatgttc tccattttca agctggaaga agtgagcagt gttggagtga 2040
 ggacctgtaa ggcaggccca ttcagagcta tgggtgctgc tgggtgcctgc caccttcaag 2100
 ttctggacct gggcatgaca tcctttcttt taatgatgcc atggcaactt agagattgca 2160
 tttttattaa agcatttcct accagcaaag caaatgttgg gaaagtattt actttttcgg 2220
 tttcaaagtgc atagaaaagt gtggcttggg cattgaaaga ggtaaaattc tctagattta 2280
 ttagtcttaa ttcaatccta cttttcgaac accaaaaatg atgcgcatca atgtatttta 2340
 tcttattttc tcaatctcct ctctctttcc tccaccata ataagagaat gttcctactc 2400
 acacttcagc tgggtcacat ccatcctcc attcatcctt ccatccatct tccatccat 2460
 tacctccatc catccttcca acatatattt attgagtacc tactgtgtgc caggggctgg 2520
 tgggacagtgc gtgacatagt ctctgccctc atagagttga ttgtctagtgc aggaagacaa 2580
 gcatttttaa aaaataaatt taaacttaca aactttgttt gtcacaagtgc gtgtttattg 2640
 caataaccgc ttggtttgca acctctttgc tcaacagaac atatgttgca agaccctccc 2700
 atgggggcac ttgagttttg gcaaggctga cagagctctg ggtgtgtgac atttctttgc 2760
 attccagctgc tcaactctgtg cttttctaca actgattgca acagactgtt gagttatgat 2820
 aacaccagtgc ggaattgctgc gaggaaccag aggcacttcc acctggctgc ggaagactat 2880
 ggtgtgcctc tgcttctgta tttccttggg ttttctgaa agtgttttta aataaagaac 2940
 aattgttaga t 2951

<210> 52
 <211> 3643
 <212> DNA
 <213> Homo sapien

<400> 52
 tgccggaaac agaataatgg cgtctcgtag cccaggcga cagcgtggag gggcgggtct 60
 gtcgattgga tgaacgcagc tgagattact cccagccact aaggacgaag aggtggggcg 120
 gtggcgctcc acgcctcgtg cgacagtggg cggggctttg ttgcctgagt aaccgtatga 180
 tgggtggtgt ggtggtgtct tcctgtctca acgataccta ttttctagtgc ctgagatcct 240
 gagacaatga gtggtatgat ccaaatgctt cacttcttcg acgttatgag cttttatttt 300

acccagggga tggatctgtt gaaatgcatg atgtaaagaa tcatcgacc tttttaaaagc 360
ggaccaaata tgataacctg cacttggaag atttatttat aggcaacaaa gtgaatgtct 420
tttctcgaca actggtatta attgactatg gggatcaata tacagctcgc cagctgggca 480
gtaggaaaga aaaaacgcta gccctaatta aaccagatgc aatatcaaag gctggagaaa 540
taattgaaat aataaacaaa gctggattta ctataaccaa actcaaaatg atgatgcttt 600
caaggaaaga agcattggat tttcatgtag atcaccagtc aagacccttt ttcaatgagc 660
tgatccagtt tattacaact ggtcctatta ttgccatgga gatTTtaaga gatgatgcta 720
tatgtgaatg gaaaagactg ctgggacctg caaactctgg agtggcacgc acagatgctt 780
ctgaaagcat tagagccctc tttggaacag atggcataag aaatgcagcg catggccctg 840
attcttttgc ttctgcgcc agagaaatgg agttgttttt tccttcaagt ggaggttggtg 900
ggccggcaaa cactgctaaa tttactaatt gtacctgttg cattgttaaa ccccatgctg 960
tcagtgaagg actgttgga aagatcctga tggctatccg agatgcaggt tttgaaatct 1020
cagctatgca gatgttcaat atggatcggg ttaatgttga ggaattctat gaagtttata 1080
aaggagtagt gaccgaatat catgacatgg tgacagaaat gtattctggc ccttgtgtag 1140
caatggagat tcaacagaat aatgctacaa agacatttcg agaattttgt ggacctgctg 1200
atcctgaaat tgcccgcat ttacgccctg gaactctcag agcaatcttt ggtaaaacta 1260
agatccagaa tgctgttcac tgtactgac tgccagagga tggcctatta gaggttcaat 1320
acttcttcaa gatcttggt aattagtgg gtggaaagta aagaagtcac aggttgggac 1380
atntagacaa gagtgaatca cacacgagga atgtgttcat tcttttattg tccgttgttt 1440
taacctgact gaatacaaga tcaacaagag cactgtactc ctggcaatta ttacatatgt 1500
tagaacatgg attttgact gtagacaaca ttaacacca gtctatggg tactgcattg 1560
ctttttataa agttcaaaat aaagatttat tttcaacaa gtgactttgg tttatttcat 1620
acattacact ttttcttgca gaaaaaata aaattgtaca actgcataaa tataaaattc 1680
ttccaccatg aaaatggta aaacattcat aaagacacag caccagcatg tgatgcctcc 1740
agaaaggaaa aaaggaaagt aaaagaaat gtataaagca aattaattag ccctgtcact 1800
agccaaagtg atgggcagat ccagatctca gacacagctt tagaacatac cggaggacaa 1860
ggcaagagat gtcaagacag cattcaatac aagcataaca caacccatca cctttttgtg 1920
tttttctgta gtgtcactta aaatttaaag ggcagttccc ccatgacaaa ccccccccc 1980
agcccccagg aaaaaataa aaataaagac ctaacaccac aggaaaagac tatggagtat 2040
gttaaaatg ctcatcgatg ggccattacc tttgaaagag ccaaaaaaca aaaaaaaaaa 2100

68

```

aaaaaaaaa attaccatgc cagtttttatt cccgttgaat atttacacct tggacagcaa 2160
accttgctca cataaagtag aaaacagata caataaaaca tggcttgaaa aatgaccaga 2220
gtatgcacct gtagtactgt aactaaata aaatacacia ggcagcaata cttaggggcc 2280
agaaacactg cttactacia gtcagttacg gaatcataat ttacagtaaa aatgggcacg 2340
tcccaaggct caatTTTTCT ttttcttttg tcatttacag tagaataaat attttgttgc 2400
tattgtctaca ctttaattta cattctaacc tattaaatgc agaaagctag tgtaaagcat 2460
atagattaag tgtaggtccc atacgtatga cagtttggtc aagactagta ggtttttgtt 2520
tttgtatctt tttttaactt attaaatggc tagtgggaaa gatttgtgct tgtgatcagc 2580
tcttaacttc aatTTTTtaca tcaaaacgtc cctgaaaacg gtctttctca ctgtacccaa 2640
tgttctcacc gtacgcctta cactctatgc gaatttcagt gtccatggta agattgggtga 2700
actgtacggc cagcaggggc tgcaggtatt tgggctgcag gagtttgcca tagtacggat 2760
aatactgcag aggaaaacca ggggagttgc ccagtccaaa atactccaca tttccaactt 2820
tatecttate ttcactctgc ttgccagtgc actgaacggg aaggacattt gggttatact 2880
tcatcactgg gtaagtctcc aaggactcat tcttgggagg cttaggtttg aagcctagaa 2940
ctcggttgag ctttataata atgcacgggt tgccctcttt gtagccataa gtttcatcat 3000
ttaatccaga gcaatttccc agccattcaa gcttgaatct gcagacctt cgctctctc 3060
gttcatgatt aaagtctcct cgttctttcg gttcactggg cacatcgcca caatcttcaa 3120
aatcatgtc atccctctgg gctgaatctt tgtacttttc caggaacctt actatgttca 3180
gtacatatgc ctcatagtc ttgggatcat taggacgaaa ggaaatttca gtcttctgga 3240
tctgaggaat ctgtgttaat cctggcgggg ccaactcggtc ctgatatgtg ggcttaaatt 3300
cactgatggt gagcagcatc acttggtatg ttccgatgaa gatgccagcc aggcagccat 3360
aaaatattac gtagaataga aggatcttaa accaactgcc accggtcctg ccagaaaact 3420
ccttcttctc tgagttccag atgaatttct tccagctgcc ctctctcttg gctttccgc 3480
gggcatggc gatggcgggt cagcagctgc cgcggtccgg aggggtgggtg gctgcggcgc 3540
gcgcctggc gtctctgcg ccgctcggt tctccggcct ggctctctgc aggcaggacg 3600
cgccgctgcc gtcgkctctc ggagtgcgag acgcgcggcg gag 3643

```

<210> 53
 <211> 1005
 <212> DNA
 <213> Homo sapien

```

<400> 53
ggccgctcgc tgacgccatg tttctctttt cggccgcgct ggtgaacagg acccgctgcc 60

```

69

```

atggggccgtg tgatccgtgg acagaggaag ggcgccgggt ctgtgttccg cgcgcacgtg 120
aagcaccgta aaggcgctgc gcgcctgcgc gccgtggatt tcgctgagcg gcacggctac 180
atcaagggca tcgtcaagga catcatccac gacccgggcc gcggcgcgcc cctcgccaag 240
gtggtcttcc gggatccgta tcggtttaag aagcggacgg agctgttcat tgccgccgag 300
ggcattcaca cgggccagtt tgtgtattgc ggcaagaagg cccagctcaa cattggcaat 360
gtgctccctg tgggcaccat gcctgagggt acaatcgtgt gctgcctgga ggagaagcct 420
ggagaccgtg gcaagctggc ccgggcatca gggaactatg ccaccgttat ctcccacaac 480
cctgagacca agaagacccg tgtgaagctg ccctccggct ccaagaaggt tatctcctca 540
gccaacagag ctgtggttgg tgtggtggct ggaggtggcc gaattgacaa acccatcttg 600
aaggctggcc gggcgtagca caaatataag gcaaagagga actgctggcc acgagtacgg 660
ggtgtggcca tgaatcctgt ggagcatcct tttggagggtg gcaaccacca gcacatcggc 720
aagccctcca ccatccgag agatgcccct gctggccgca aagtgggtct cattgctgcc 780
cgccggactg gacgtctccg gggaaccaag actgtgcagg agaayatgtg ggckcmaagt 840
gggtttgccc aaaaaaacac aacaacgcaa acacaaaggc aaacatatag aaaaaaggg 900
ggctaccggg gccgacactc tagaggtaat ataatagcgg cagaggacag gggagggtag 960
ggcaacactt tattctgtat atttggccgc tctaaaatca cactc 1005

```

```

<210> 54
<211> 3273
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (3246)..(3246)
<223> n=a, c, g or t

```

```

<400> 54
gccatggtaa gcgcggacgc catggttaagc gcggacgccca tggatatgta agcgcggacg 60
ccatggtaag cgcgacgcc atggttaagcg cggacccat ggtaagcgcg gacgccatgg 120
taagcgcgga cgccatggta agcgcggacg ccatggtaag cgcgacgcc atggttaagcg 180
cggacgccat ggtaagcgcg gacgccatgc acacggaccc tgactactcg gctgcctatg 240
tcgtcataga aactgatgca gaagatggaa tcaaggggtg tggaattacc ttcactctgg 300
gaaaaggcac tgaagttggt gtctgtgctg tgaatgccct cgcccacat gtgctcaaca 360
aggacctcaa ggacattggt ggtgacttca gaggttcta taggcagctc acaagtgatg 420
ggcagctcag atggattggt ccagaaaagg gcgtgggtgca cctggcgaca gcggccgtcc 480

```

70

taaacgcggt gtgggacttg tgggccaagc aggagggaaa gcctgtctgg aagttacttg	540
tggacatgga tcccaggatg ctggtatcct gcatagattt caggtagatc actgatgtcc	600
tgactgagga ggatgcccta gaaatactgc agaaagggtca aattggtaaa aaagaaagag	660
agaagcaaat gctggcaciaa ggataccctg cttacacgac atcgtgcgcc tggctggggg	720
actcagatga cacgttgaag cagctctgtg cccaggcgct gaaggatggc tggaccagg	780
ttaaagtaaa ggtgggtgct gatctccagg atgacatgcg aagatgccaa atcatccgag	840
acatgattgg accggaaaag actttgatga tggatgccaa ccagcgctgg gatgtgcctg	900
aggcggtgga gtggatgtcc aagctggcca agttcaagcc attgtggatt gaggagccaa	960
cctcccctga tgacattctg gggcacgcca ccatttccaa ggactgggtc ccattaggaa	1020
ttggcattgc cacaggagaa cagtgccaca atagagtgat atttaagcaa ctctacagg	1080
cgaaggccct gcagttcctc cagattgaca gttgcagact gggcagtgct aatgagaacc	1140
tctcagtatt gctgatggcc aaaaagtttg aaattcctgt ttgccccat gctggtggag	1200
ttggcctctg tgaactgggtg cagcacctga ttatatattga ctacatatca gtttctgcaa	1260
gccttgaaaa tagggtgtgt gagtatgttg accacctgca tgagcatttc aagtatcccg	1320
tgatgatcca gcgggcttcc tacatgctc ccaaggatcc cggctactca acagaaatga	1380
aggaggaatc tgtaaagaaa caccagtatc cagatgggtga agtttggaag aaactccttc	1440
ctgctcaaga aaattaagtg ctacagccca acaacttttt tctttctgaa gtgaaagggc	1500
ttaaaatttc ttggaaatag ttttcaaaaa atggatttaa aaaatcctac cgatcaagat	1560
gagttcagct agaagtcata ccacctcag gaatcagcta aagcaaaaag aacttttacc	1620
tcggcatcca gcccaacccc taaagactga caatattcct cgagctcctt tgaaagcacc	1680
ctaaacagcc atttccattt taatagttgg atgctgattg tacccttcaa tctgaaagtc	1740
ttcagctttg aagtcatcaa ttttctcaac ttttcgaaga atcctgagct ttgggaaagg	1800
tctgggttct cgctgaagct gaattttcag tggctcgatg tgattcagg	1860
atctcccaaa gtgtgtataa agtcacctg ctataaaccc gggggagaaa gcagaacagt	1920
atgttagttt caattcttta aaacatcatt taaaaacatt agaatatgca gacaccgcaa	1980
ggcttttttt aaaaaataa tttagtgtag cttttccatt tttttgtagc aacagcatct	2040
tggtatgttg cccaggctgg tattgaactc cagacctcaa gcaattgctc ctgtctcagt	2100
ctcccaaagt gctgggatta caggcatgag ccaccatacc caacctcagc atagcttttg	2160
agaaaatcca tagaagctgt atcacaaca acctgtatag atctgttagt gcgtatacca	2220
cagggccaga aaaccttcca gaagaggaag gtttcaaagt aaaagctgg	2280
cttacacata tcaaatttaa aagctaata gagactaaac tctgcaattt gttttcccat	2340

attaaagaac tgaagagctc agtgtggttag gctggcaagt cacccttccc gagacagccc 2400
 acccttcaggc ccgtgatgtg cgcaatcatg tacgtgagca gggcgtagct ggcgatgttg 2460
 aaaggcacac cgaggcccat gtctcccgat ctctggtaca gctggcagga cagctcactg 2520
 ttcaccacat agaactggca gagggcatgg catggaggca ggcccatcag aggaagatct 2580
 cttggattcc aagcgacat gatgattctt ctgtcgtcag ggttggtttt gatggtgtca 2640
 atcactcttt gcagttggtc aactccctgt cctgaataat ctgattccat atctctgtat 2700
 tctgccccaa aatgcctcca ctggaagcca taaactgggc ccaagtcccc ttcttctctg 2760
 gtggagaatc ccaggctgtc caaaaagtct cgggatccat tggcatccca gattttcact 2820
 cccttggaag acagctcttt agcatttgtg gatcccttga taaaccacag caactcctcc 2880
 aaaacaccct tccagaacac acgtttggtt gtcagcagag ggaattcatc tctcaggctg 2940
 tagcgcgcct gcatgccgaa taccgacagg gtgccggtgc ccgtgcggtc gtccttctctg 3000
 acgccgcagc ggaggatgtg ttggatctgc ccagggtact gcagctcccc gtgcggcgga 3060
 cgcggctcgg cgtcccgtc ctgtgcggcg gggggcaagg gccggcgcgg cagctccgag 3120
 ccggccacag gcatggcgcg gcggggcggg gacggaggca ggcgaagtgg cgcggcgggga 3180
 cggaggcagg ccaagtggcg cggcgggacg gaggcaggcc aagtggcgcg gtggcaggac 3240
 cctcgnccga ttaaccgagc ctctactag cca 3273

<210> 55
 <211> 2967
 <212> DNA
 <213> Homo sapien

<400> 55
 gccatggtaa gcgcggacgc catggttaagc gcggacgcca tggatatgta agcgcggacg 60
 ccatggttaag cgcggacgcc atggttaagcg cggacgccat ggtaagcgcg gacgccatgg 120
 taagcgcgga cgccatggtta agcgcggacg ccatggttaag gcgcggacgcc atggttaagcg 180
 cggacgccat ggtaagcgcg gacgccatgc acacggaccc tgactactcg gctgcctatg 240
 tcgtcataga aactgatgca gaagatggaa tcaaggggtg tggaattacc ttcactctgg 300
 gaaaaggcac tgaagttggt gtctgtgctg tgaatgccct cgcccacccat gtgctcaaca 360
 aggacctcaa ggacattggt ggtgacttca gaggcttcta taggcagctc acaagtgatg 420
 ggcagctcag atggattggt ccagaaaagg gcgtggtgca cctggcgaca gcggccgtcc 480
 taaacgcggt gtgggacttg tgggccaagc aggagggaaa gcctgtctgg aagttacttg 540
 tggacatgga tcccaggatg ctggtatcct gcatagattt caggtacatc actgatgtcc 600
 tgactgagga ggatgcccta gaaatactgc agaaaggcca aattggtaaa aaagaaagag 660

agaagcaaat gctggcacia ggataccctg cttacacgac atcgtgcgcc tggctggggt 720
actcagatga cacgttgaag cagctctgtg cccaggcgct gaaggatggc tggaccaggt 780
ttaaagtaaa ggtgggtgct gatctccagg atgacatgcg aagatgccaa atcatccgag 840
acatgattgg accggaagaa actttgatga tggatgccaa ccagcgctgg gatgtgcctg 900
aggcgggtgga gtggatgtcc aagctggcca agttcaagcc atttgtgatt gaggagccaa 960
cctcccctga tgacattctg gggcacgcca ccatttccaa ggcaactggc ccattaggaa 1020
ttggcattgc cacaggagaa cagtgcaca atagagtgat atttaagcaa ctccctacagg 1080
cgaaggccct gcagttcctc cagattgaca gttgcagact gggcagtgtc aatgagaacc 1140
tctcagatatt gctgatggcc aaaaagtgtg aaattcctgt ttgcccccat gctggtggag 1200
ttggcctctg tgaactgggtg cagcacctga ttatatgtga ctacatatca gtttctgcaa 1260
gccttgaaaa tagggtgtgt gagtatgtt accacctgca tgagcatttc aagtatcccg 1320
tgatgatcca gcgggcttcc tacatgcctc ccaaggatcc cggctactca acagaaatga 1380
aggaggaatc tgtaaagaaa caccagtatc cagatgggtga agtttggaag aaactccttc 1440
ctgctcaaga aaattaagtg ctgagcccca acaacttttt tctttctgaa gtgaaagggc 1500
ttaaaatttc ttgaaatag ttttcaaaa atggatttaa aaaatcctac cgatcaagat 1560
gagttcagct agaagtcata ccaccctcag gaatcagcta aagcaaaaag aacttttacc 1620
tcggcatcca gcccaacccc taaagactga caatatecctt cgagctcctt tgaaagcacc 1680
ctaaacagcc atttccattt taatagttgg atgcggattg tacccttcaa tctgaaagtc 1740
ttcagctttg aagtcataaa ttttctcaac ttttcaaga atcctgagct ttgggaaagg 1800
tctgggttct cgctgaagct aaaaacaaaa taaggccatt attttgccat aattgtacga 1860
cctgttgtaa ttgctcctca tgtccgtgaa acaagtacac aggatgtgat caacaaagtt 1920
ctattttaca ggagtatgat cctgtcgata ccttgccgta ggttatgtaa catgattgga 1980
gcgcaaccag ctgttctctt gcacagatcg agagtgaagg gtattttgtg acattacaca 2040
gcacagagg cctggtgcct catcaggtaa ctgaagagct cagtgtggta ggctggcaag 2100
tcacccttcc cgagacagcc caccttcagg cccgtgatgt gcgcaatcat gtacgtgagc 2160
agggcgtagc tggcgatgtt gaaaggcaca ccgaggccca tgtctccga tctctggtac 2220
agctggcagg acagctcact gttcaccaca tagaactggc agagggcatt gcattggaggc 2280
agcgccatca gaggaagatc tcttgatttc caagcgaca tgatgattct tctgtcgta 2340
gggttggttt tgatgggtgc aatcactctt tgcagttggt caactcctg tctgaataa 2400
tctgattcca tatctctgta ttctgcccc aatgcctcc actggaagcc ataaactggg 2460

73

```

cccaagtccc cttcttctct ggtggagaat ccagggtgt ccaaaaagtc tcgggatcca 2520
ttggcatccc agattttcac tcccttgga gacagctctt tagcatttgt ggatcccttg 2580
ataaaccaca gcaactcctc caaaacaccc ttccagaaca cacgtttggt tgtcagcaga 2640
gggaattcat ctctcaggct gtagcgcgcc tgcattgcca ataccgacag ggtgccggtg 2700
cccgtgcggt cgtccttcct gacgccgcag cggaggatgt gttggatctg ccccaggtag 2760
tgcagctccc cgtgcgggcg acgcggctcg gcgtcccgct cctgtgcggc ggggggcaag 2820
ggccggcgcg gcagctccga gccggccaca ggcatggcg gcggggcggg ggacggaggc 2880
aggcgaagtg gcgcggcggg acggaggcag gccaaagtggc gcggggcagg accccttccg 2940
cgcgcttttt ccgcgggcgg tcccgcg 2967

```

```

<210> 56
<211> 2761
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (2734)..(2734)
<223> n=a, c, g or t

```

```

<400> 56
gccatggtaa gcgcggacgc catggttaagc gcggacgcca tggatatgta agcgcgacg 60
ccatggtaag gcgcggacgc atggttaagc gcggacgcca ggtaagcgcg gacgccatgg 120
taagcgcgga cgccatggta agcgcgacgc ccatggtaag gcggacgcca atggttaagc 180
cggacgccaat ggtaagcgcg gacgccatgc acacggaccc tgactactcg gctgcctatg 240
tcgtcataga aactgatgca gaagatggaa tcaaggggtg tgggaattacc ttcactctgg 300
gaaaaggcac tgaagttggt gtctgtgctg tgaatgccct cgcccacat gtgctcaaca 360
aggacctcaa ggacattggt ggtgacttca gaggcttcta taggcagctc acaagtgatg 420
ggcagctcag atggattggt ccagaaaagg gcgtggtgca cctggcgaca gcggccgtcc 480
taaacgcggt gtgggacttg tgggccaagc aggagggaaa gcctgtctgg aagttacttg 540
tggaatgga tcccaggatg ctggtatcct gcatagattt caggtagatc actgatgtcc 600
tgactgagga ggatgcccta gaaatactgc agaaagggtc aattggtaaa aaagaaagag 660
agaagcaaat gctggcacia ggataccctg cttacacgac atcgtgcgcc tggctggggg 720
actcagatga cacgttgaag cagctctgtg cccaggcgct gaaggatggc tggaccagg 780
ttaaagtaaa ggtgggtgct gatctccagg atgacatgcg aagatgcaa atcatccgag 840
acatgattgg accgaaaaag actttgatga tggatgcaa ccagcgctgg gatgtgcctg 900

```

aggcgggtgga gtggatgtcc aagctggcca agttcaagcc attgtggatt gaggagccaa 960
cctcccctga tgacattctg gggcacgcca ccatttccaa ggcactggtc ccattaggaa 1020
ttggcattgc cacaggagaa cagtgccaca atagagtgat atttaagcaa ctctacagg 1080
cgaaggccct gcagttcctc cagattgaca gttgcagact gggcagtgtc aatgagaacc 1140
tctcagtatt gctgatggcc aaaaagtttg aaattcctgt ttgccccat gctggtggag 1200
ttggcctctg tgaactggg cagcacctga ttatatttga ctacatatca gtttctgcaa 1260
gccttgaaaa taggggtgtg gagtatgtg accacctgca tgagcatttc aagtatcccg 1320
tgatgatcca gcgggcttcc tacatgcctc ccaaggatcc cggctactca acagaaatga 1380
aggaggaatc tgtaaagaaa caccagtatc cagatggtga agtttggaag aaactccttc 1440
ctgctcaaga aaattaagtg ctacgcccc acaacttttt tctttctgaa gtgaaagggc 1500
ttaaatttct ttggaaatag ttttcaaaaa atggatttaa aaaatcctac cgatcaagat 1560
gagttcagct agaagtcata ccacctcag gaatcagcta aagcaaaaag aacttttacc 1620
tcggcatcca gcccaacccc taaagactga caatctctt cgagctcctt tgaaagcacc 1680
ctaaacagcc atttccattt taatagttgg atgcggttg tacccttcaa tctgaaagtc 1740
ttcagctttg aagtcatcaa ttttctcaac ttttcgaaga atcctgagct ttgggaaagg 1800
tctgggttct cgctgaagct gaattttcag tggctcgatg tgattcaggt aaatatgtgc 1860
atctcccaaa gtgtgtataa agtcacctg cttcaggccc gtgatgtgcg caatcatgta 1920
cgtgagcagg gcgtagctgg cgatgttgaa aggcacaccg aggcccatgt ctcccgatct 1980
ctggtacagc tggcaggaca gtcactgtt caccacatag aactggcaga gggcatggca 2040
tggaggcagc gccatcagag gaagatctct tggattccaa gcgcacatga tgattcttct 2100
gtcgtcaggg ttggttttga tgggtgcaat cactctttgc agttggtcaa ctccctgtcc 2160
tgaataatct gattccatat ctctgtattc tgccccaaaa tgctccact ggaagccata 2220
aactgggccc aagtccccct cttctctggt ggagaatccc aggctgtcca aaaagtctcg 2280
ggatccattg gcatcccaga ttttactcc cttggaagac agctctttag catttgtgga 2340
tcccttgata aaccacagca actcctccaa aacacccttc cagaacacac gtttggttgt 2400
cagcagaggg aattcatctc tcaggctgta gcgcgcctgc atgccgaata ccgacaggg 2460
gccggtgccc gtgcggtcgt cttctctgac gccgcagcg aggatgtgtt ggatctgccc 2520
caggtactgc agctccccgt gcggcgagcg cggtcggcg tcccgtcct gtgcggcggg 2580
gggcaagggc cggcgcgga gctccgagcc ggccacaggc atggcgcggc gggcggggga 2640
cggaggcagg cgaagtggcg cggcgggagc gaggcaggcc aagtggcgcg gcgggacgga 2700
ggcaggccaa gtggcgcggt ggcaggaccc tcgnccgatt aaccgagcct cgtactagcc 2760

a

2761

<210> 57
<211> 3485
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (3458)..(3458)
<223> n=a, c, g or t

<400> 57
gccatggtaa gcgcggacgc catggtaagc gcggacgcca tggatatggt agcgcggacg 60
ccatggtaag gcgcggacgcc atggtaagcg cggacgccat ggtaagcgcg gacgccatgg 120
taagcgcgga cgccatggta agcgcggacg ccatggtaag gcgcggacgcc atggtaagcg 180
cggacgccat ggtaagcgcg gacgccatgc acacggaccc tgactactcg gctgcctatg 240
tcgtcataga aactgatgca gaagatggaa tcaaggggtg tggaattacc ttactctggt 300
gaaaaggcac tgaagttggt gtctgtgctg tgaatgccct cgcccacccat gtgctcaaca 360
aggacctcaa ggacattggt ggtgacttca gaggcttcta taggcagctc acaagtgatg 420
ggcagctcag atggattggt ccagaaaagg gcgtgggtgca cctggcgaca gcggccgtcc 480
taaacgcggt gtgggacttg tgggccaagc aggagggaaa gcctgtctgg aagttacttg 540
tggacatgga tcccaggatg ctggtatcct gcatagattt caggtagatc actgatgtcc 600
tgactgagga ggatgcccta gaaatactgc agaaagggtc aattggtaaa aaagaaagag 660
gtgggttgta agaaaatttt cttcattggt tttgctaaca ttgtccactt ttgagtcccc 720
ctgtcctttt ggggtacaca ttgtcttccc aaatgccctg tgctgagcag ctaggccctc 780
aaatcaacat tcaagtctgc atggtgaagc ctgctgggta tgacctctga ctgcagagtt 840
tgcttcagcc actgctgaaa ggaagtttgg ctttaggatt acactgtagg gagagccctg 900
ggggagcagg gcagtcctg agagtatcct gatcacctgg gtttgacatc ctagtaattt 960
gtggctgggt gtgtgtgtgc agggccggat caggagaaca gctggactct ccaggggaaa 1020
cagcttagct acaggcactt ccaattccga agggccctgg aaagtgcaaa atgttgacgg 1080
cgctgtgttt tcacagagaa gcaaatgctg gcacaaggat accctgctta cacgacatcg 1140
tgcgcctggc tggggtactc agatgacacg ttgaagcagc tctgtgcca ggcgctgaag 1200
gatggctgga ccaggtgagt gtgatgatgg acctgacttt ccagttggc ggcaggagag 1260
actcaggcag tttaaagtaa aggtgggtgc tgatctccag gatgacatgc gaagatgcca 1320
aatcatccga gacatgattg gaccggaaaa gactttgatg atggatgcca accagcgctg 1380

ggatgtgcct gaggcggtgg agtggatgtc caagctggcc aagttcaagc cattgtggat 1440
tgaggagcca acctcccctg atgacattct ggggcacgcc accatttcca aggcactggt 1500
cccattagga attggcattg ccacaggaga acagtgccac aatagagtga tatttaagca 1560
actcctacag gcgaaggccc tgcagttcct ccagattgac agttgcagac tgggcagtgt 1620
caatgagaac ctctcagtat tgctgatggc caaaaagttt gaaattcctg tttgccccca 1680
tgctggtgga gttggcctct gtgaactggt gcagcacctg attatatttg actacatata 1740
agtttctgca agccttgaaa ataggggtgtg tgagtatggt gaccacctgc atgagcattt 1800
caagtatccc gtgatgatcc agcgggcttc ctacatgcct cccaaggatc ccggctactc 1860
aacagaaatg aaggaggaat ctgtaaagaa acaccagtat ccagatgggtg aagtttgga 1920
gaaactcctt cctgctcaag aaaattaagt gctcagcccc aacaactttt ttctttctga 1980
agtgaagggt cttaaaattt cttggaaata gttttacaaa aatggattta aaaaatccta 2040
ccgatcaaga tgagttcagc tagaagtcac accaccctca ggaatcagct aaagcaaaaa 2100
gaacttttac ctcgcatcc agcccaaccc cttaaagactg acaatatact tcgagctcct 2160
ttgaaagcac cctaaacagc catttccatt ttaatagttg gatgcggatt gtacccttca 2220
atctgaaagt cttcagcttt gaagtcacac attttctcaa cttttcgaag aatcctgagc 2280
tttgggaaag gtctgggttc tcgctgaagc taaaaacaaa ataaggccat tattttgcca 2340
taattgtacg acctgttgta attgctcttc atgtccgtga aacaagtaca caggatgtga 2400
tcaacaaagt tctattttac aggagtatga tcctgtcgat accttgccgt aggttatgta 2460
acatgattgg agcgcaacca gctgttctct tgcacagatc gagagtgagg ggtattttgt 2520
gacattacac agcatcagga gcctgggtgcc tcatcaggta actgaagagc tcagtgtggt 2580
aggctggcaa gtcacccttc ccgagacagc ccaccttcag gcccgatgatg tgcgcaatca 2640
tgtacgtgag cagggcgtag ctggcgatgt tgaaaggcac accgaggccc atgtctcccc 2700
atctctggta cagctggcag gacagctcac tgttcaccac atagaactgg cagagggcat 2760
ggcatggagg cagcgccatc agaggaagat ctcttggtt ccaagcgcac atgatgatcc 2820
ttctgtcgtc aggggttggt ttgatgggt caatcactct ttgcagttgg tcaactcctt 2880
gtcctgaata atctgattcc atatctctgt attctgcccc aaaatgcctc cactggaagc 2940
cataaactgg gcccaagtcc cttcttcttc tgggtggagaa tcccaggctg tccaaaaagt 3000
ctcgggatcc attggcatcc cagattttca ctcccttgga agacagctct ttagcatttg 3060
tggatccctt gataaaccac agcaactcct ccaaaacacc cttccagaac acacgtttgg 3120
ttgtcagcag agggaattca tctctcaggc tgtagcgcgc ctgcatgccg aataccgaca 3180

77

ggggtgccggt gcccgtagcg tgcgtcctcc tgacgccgca gcggaggatg tgttgatct 3240
 gccccaggta ctgcagctcc ccgtgcggcg gacgcggctc ggcgccccgc tcctgtgcgg 3300
 cgggggggcaa ggcccgggcg ggcagctccg agccggccac aggcattggc cggcggggcg 3360
 gggacggagg caggcgaagt ggcgcggcg gacggaggca ggccaagtgg cgcggcggga 3420
 cggaggcagg ccaagtggcg cggtaggcagg accctcgncg gattaaccga gcctcgtact 3480
 agcca 3485

<210> 58
 <211> 4112
 <212> DNA
 <213> Homo sapien
 <220>
 <221> misc_feature
 <222> (4085)..(4085)
 <223> n=a, c, g or t

<400> 58
 gccatggtaa gcgcggacgc catggttaag gcggacgcca tggtaggta agcgcggacg 60
 ccatggttaag gcgcggacgc atggttaagc cggacgccat ggtaagcgcg gacgccatgg 120
 taagcgcgga cgcattggtta agcgcggacg ccatggttaag gcgcggacgc atggttaagc 180
 cggacgccat ggtaagcgcg gacgccatgc acacggaccc tgactactcg gctgcctatg 240
 tcgtcataga aactgatgca gaagatggaa tcaaggggtg tggaattacc ttcactctgg 300
 gaaaaggcac tgaagtgtgt gtctgtgctg tgaatgccct cccccaccat gtgtcaaca 360
 aggacctcaa ggacattgtt ggtgacttca gaggcttcta taggcagctc acaagtgatg 420
 ggcagctcag atggattggt ccagaaaagg gcgtggtgca cctggcgaca gcggccgtcc 480
 taaacgcggt gtgggacttg tgggccaagc aggagggaaa gcctgtctgg aagttacttg 540
 tggacatgga tcccaggatg ctggtatcct gcatagattt caggtagatc actgatgtcc 600
 tgactgagga ggatgcccta gaaatactgc agaaagggtca aattggtaaa aaagaaagag 660
 agaagcaaat gctggcaca gataaccctg cttacacgac atcgtgcgcc tggctgggg 720
 actcagatga cacgttgaag cagctctgtg cccaggcgct gaaggatggc tggaccagg 780
 ttaaagtaaa ggtgggtgct gatctccagg atgacatgcg aagatgccaa atcatccgag 840
 acatgattgg accggaaaag actttgatga tggatgccaa ccagcgtgg gatgtgcctg 900
 aggcgggtgga gtggatgtcc aagctggcca agttcaagcc attgtggatt gaggagccaa 960
 cctcccctga tgacattctg gggcacgcca ccatttccaa ggcaactggtc ccattaggaa 1020
 ttggcattgc cacaggagaa caggtagatg acgcccccaa caggtagatg acgtcccctt 1080

ggggtcagta cacgctgacc agtgaccgag gacacagttg tgtgttaggc tccatcacct 1140
gctgtacttt gagttgggaa attttcatca tcttagaaac tgggtcattt tatcagagtc 1200
tagagtcaga tatagaaaaa gtttgtggct atttctccaa tttatatgac taaggtcagg 1260
tatctttttc aaagtgtcta attgaaattg aaaaggcagc aatttaaagt tgctattgca 1320
agggcagaaa atggtcttaa gaaagccagc tttcaaattg aataaacatg actgcgttca 1380
ctttttgagc ttataaatga agcccgagtg cctgccaaaa cctgctgcag tcagcccacg 1440
agcagagcag cgtgaggagc tgattctcag ttttcccggc aaaaggagca atactgctct 1500
gccgtgggtc cgtgttgtca tctgtgccac ctgctcatca ctgtcacctg atttcacct 1560
gatgcttcat ctcccactta tcagtcgctg tgacagtcac tccctcataa atggcgagcc 1620
agtgtgattt tgacctgact cacactgttg cattagcaga tttgtaaaga agtgagcaca 1680
aggtccttgc ccacgctata aaagctcgcc tcatgccagc cgagaacaaa gaagaaatac 1740
agtctgggct tcctgacggc cactgatgaa taattattgg catagagtgg ctgcgttgcc 1800
aggtttagag atcctgaagg ccaaggctga ctcttctgtt ggtgttattt tcaattctat 1860
ttccagtgcc acaatagagt gatatttaag caactcctac aggcgaaggc cctgcagttc 1920
ctccagattg acagttgcag actgggcagt gtcaatgaga acctctcagt attgctgatg 1980
gccaaaaagt ttgaaagtaa gcgtgctgca ggggctgcag accagacctt catttcccca 2040
ctaatacagac acctcccttg atggtttgca attcacatgc atgggagtct gtagtttgcc 2100
atttcgattt ttttctaact ctcathtagc tttaatccgg gaatttttga ttttcatctt 2160
ggaatttcct tcctaaatat taataaatga tttaatcacc tgtgggcaat aagaaaaacc 2220
agaaagtccc ctttcacccc ttctctccc tgccctactc ttggtattaa atagaaacga 2280
tttctttta gttctgttt gccccatgc tgggtggagt ggctctctgt aactggtgca 2340
gcacctgatt atatttgact acatatcagt ttctgcaagc cttgaaaata ggggtgtgta 2400
gtatgttgac cacctgcagt agcatttcaa gtatcccgtg atgatccagc gggcttccta 2460
catgcctccc aaggatcccg gctactcaac agaaatgaag gaggaatctg taaagaaaca 2520
ccagtatcca gatggtgaag tttggaagaa actccttcct gctcaagaaa attaatgtgt 2580
cagccccaac aacttttttc tttctgaagt gaaagggctt aaaatttctt ggaaatagtt 2640
ttacaaaaat ggatttaaaa aatcctaccg atcaagatga gttcagctag aagtcatacc 2700
acctcagga atcagctaaa gcaaaaagaa cttttacctc ggcatccagc ccaacccta 2760
aagactgaca atatccttcg agctcctttg aaagcacctt aaacagccat ttccatttta 2820
atagttggat gcggattgta cccttcaatc tgaaagtctt cagctttgaa gtcatcaatt 2880
ttctcaactt ttcgaagaat cctgagcttt gggaaaggct tgggttctcg ctgaagctaa 2940

```

aaacaaaata aggccattat ttgccataa ttgtacgacc tgttgtaatt gtcctcatg 3000
tccgtgaaac aagtacacag gatgtgatca acaaagttct attttacagg agtatgatcc 3060
tgtcgatacc ttgccgtagg ttatgtaaca tgattggagc gcaaccagct gttctcttgc 3120
acagatcgag agtgaggggt attttgtgac attacacagc atcaggagcc tgggtgcctca 3180
tcaggtaact gaagagctca gtgtggtagg ctggcaagtc acccttcccg agacagccca 3240
ccttcaggcc cgtgatgtgc gcaatcatgt acgtgagcag ggcgtagctg gcgatgttga 3300
aaggcacacc gagggccatg tctcccgatc tctggtagc ctggcaggac agctcactgt 3360
tcaccacata gaactggcag agggcatggc atggaggcag cgccatcaga ggaagatctc 3420
ttggattcca agcgcacatg atgattcttc tgcgctcagg gttgggtttg atgggtgcaa 3480
tactctttg cagttggtca actccctgtc ctgaataatc tgattccata tctctgtatt 3540
ctgccccaaa atgcctccac tggagccat aaactgggcc caagtcccct tcttctctgg 3600
tggagaatcc caggctgtcc aaaaagtctc gggatccatt ggcacccag attttctc 3660
ccttgaaga cagctcttta gcatttgtgg atcccttgat aaaccacagc aactcctcca 3720
aaacaccctt ccagaacaca cgtttggttg tcagcagagg gaattcatct ctcaggctgt 3780
agcgcgctg catgccgaat accgacaggg tgccggtgcc cgtgcggtcg tccttctga 3840
cgccgcagcg gaggatgtgt tggatctgcc ccaggtagc cagctccccg tgcggcggac 3900
gcggctcggc gtcccgtcc tgtgcggcgg gggcaaggg ccggcgcggc agctccgagc 3960
cgccacagg catggcgcg cgggcggggg acggaggcag gcgaagtggc gcggcgggac 4020
ggaggcaggc caagtggcgc ggcgggacgg aggcaggcca agtggcgcg tggcaggacc 4080
ctcgnccgat taaccgagcc tcgtactagc ca 4112

```

```

<210> 59
<211> 2636
<212> DNA
<213> Homo sapien

```

```

<400> 59
gccatggtaa gcgcggacgc catggttaagc gcggacgcca tggtaggta agcgcggacg 60
ccatggttaag gcgcggacgcc atggttaagc cgacgccat ggtaagcgcg gacgccatgg 120
taagcgcgga cgccatggta agcgcggacg ccatggttaag gcgcggacgcc atggttaagc 180
cggacgccat ggtaagcgcg gacgccatgc acacggaccc tgactactcg gctgcctatg 240
tcgtcataga aactgatgca gaagatggaa tcaaggggtg tgggaattacc ttcactctgg 300
gaaaaggcac tgaagttggt gtctgtgctg tgaatgccct cgcccacat gtgtcaaca 360
aggacctcaa ggacattggt ggtgacttca gaggccttcta taggcagctc acaagtgatg 420

```

ggcagctcag atggattggt ccagaaaagg gcggtggtgca cctggcgaca gcggccgtcc 480
taaacgcggt gtgggacttg tgggccaagc aggagggaaa gcctgtcttg aagttacttg 540
tggacatgga tcccaggatg ctggtatcct gcatagattt caggtacatc actgatgtcc 600
tgactgagga ggatgcccta gaaatactgc agaaagggtca aattggtaaa aaagaaagag 660
agaagcaaat gctggcacia ggataccctg cttacacgac atcgtgcgcc tggctggggt 720
actcagatga cacgttgaag cagctctgtg cccaggcgct gaaggatggc tggaccaggt 780
ttaaagtaaa ggtgggtgct gatctccagg atgacatgcg aagatgccaa atcatccgag 840
acatgattgg accggaaaag actttgatga tggatgccaa ccagcgctgg gatgtgcctg 900
aggcggtgga gtggatgtcc aagctggcca agttcaagcc attgtggatt gaggagccaa 960
cctcccctga tgacattctg gggcacgcca ccatttccaa ggcaactggc ccattaggaa 1020
ttggcattgc cacaggagaa cagtgccaca atagagtgat atttaagcaa ctccacagg 1080
cgaaggccct gcagttctc cagattgaca gttgcagact gggcagtgct aatgagaacc 1140
tctcagtatt gctgatggcc aaaaagtttg aaattcctgt ttgccccat gctggtggag 1200
ttggcctctg tgaactggtg cagcacctga ttatatttga ctacatatca gtttctgcaa 1260
gccttgaaaa tagggtgtgt gagtatgttg accacctgca tgagcatttc aagtatcccg 1320
tgatgatcca gcgggcttcc tacatgcctc ccaaggatcc cggctactca acagaaatga 1380
aggaggaatc tgtaaagaaa caccagtatc cagatgggtga agtttgaag aaactccttc 1440
ctgctcaaga aaattaagtg ctacgcccc acaacttttt tctttctgaa gtgaaagggc 1500
ttaaatttct ttggaatag ttttataaaa atggatttaa aaaatcctac cgatcaagat 1560
gagttcagct agaagtcata ccacctcag gaatcagcta aagcaaaaag aacttttacc 1620
tcggcatcca gcccaacccc taaagactga caatatcctt cgagctcctt tgaaagcacc 1680
ctaaacagcc atttccattt taatagttgg atgcggattg tacccttcaa tctgaaagtc 1740
ttcagctttg aagtcatcaa ttttctcaac ttttcgaaga atcctgagct ttgggaaagg 1800
tctgggttct cgctgaagct aaaaacaaaa taaggccatt attttgccat aattgtacga 1860
cctgttgtaa ttgctcctca tgtccgtgaa acaagtacac aggatgtgat caacaaagtt 1920
ctattttaca ggagtatgat cctgtcgata ccttgccgta gggtatgtaa catgattgga 1980
gcgcaaccag ctgttctctt gcacagatcg agagtgaggg gtattttgtg acattacaca 2040
gcacagagag cctggtgect catcaggtgt aagttcttat aaccactctt ggcaaattta 2100
ttaaagacag gaacacagtc aatctgtaac tcatagtagc tctacgttta cttgaattcc 2160
acaatcccta acccatctgt ccctggcaga aagaaggaaa gatgacatgc atggacagtg 2220

81

aacagaaagg gatgaaagcc aggattcctg ggatgaacag acagtggcaa ttaggatgtg 2280
aagacaggtc acaacctatt actatgtcta aaaacgacca gagcagagag ccagagagaa 2340
taagcctgaa gtcacctcca ctcaaaagca gccaaactcc ctcaaaggag taacttttaa 2400
aacctggatc taacctggaa ggggctaaaa agtgtctggt tctgagtttt tttccttaag 2460
gctcatgaag cagatgaact tacattttta ttgccatttc atatcaattg ttggctgcta 2520
taacttaggg atttcaacag acttttgaag tttggaccta aatattgtac ttaatgtaaa 2580
ttaacaaaaa atatttatgg ccagggtggt ggcttatgcc tgtaattcca gaattt 2636

<210> 60
<211> 2764
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (2737)..(2737)
<223> n=a, c, g or t

<400> 60
gccatggtaa gcgcgacgc catggttaagc gcggacgcca tggatatgta agcgcggacg 60
ccatggtaag cgcggacgcc atggtaagcg cggacgccat ggtaagcgcg gacgccatgg 120
taagcgcgga cgccatggtg agcgcggacg ccatggtaag cgcggacgcc atggtaagcg 180
cggacgccat ggtaagcgcg gacgccatgc acacggaccc tgactactcg gctgcctatg 240
tcgtcataga aactgatgca gaagatggaa tcaaggggtg tggaattacc ttcactctgg 300
gaaaaggcac tgaagttggt gtctgtgctg tgaatgcctt cgcccaccat gtgctcaaca 360
aggacctcaa ggacattggt ggtgacttca gaggtctcta taggcagctc acaagtgatg 420
ggcagctcag atggattggt ccagaaaagg gcgtggtgca cctggcgaca gcggccgtcc 480
taaacgcggg gtgggacttg tgggccaagc aggagggaaa gcctgtctgg aagttacttg 540
tggacatgga tcccaggatg ctggtatcct gcatagattt caggtaacatc actgatgtcc 600
tgactgagga ggatgcccta gaaatactgc agaaagggtc aattggtaaa aaagaaagag 660
agaagcaaat gctggcacia ggataccctg cttacacgac atcgtgcgcc tggctggggg 720
actcagatga cacgttgaag cagctctgtg cccaggcgct gaaggatggc tggaccaggt 780
ttaaagtaaa ggtgggtgct gatctccagg atgacatgcy aagatgccaa atcatccgag 840
acatgattgg accggaaaag actttgatga tggatgccaa ccagcgtggt gatgtgcctg 900
aggcgggtga gtggatgtcc aagctggcca agttcaagcc attgtggatt gaggagccaa 960
cctcccttga tgacattctg gggcacgcca ccatttccaa ggcactgggt ccattaggaa 1020

ttggcattgc cacaggagaa caggggtgtgt gagtatgttg accacctgca tgagcatttc	1080
aagtatcccc tgatgatcca gcgggcttcc tacatgcctc ccaaggatcc cggctactca	1140
acagaaatga aggaggaatc tgtaaagaaa caccagtatc cagatgggtga agtttggaag	1200
aaactccttc ctgctcaaga aaattaagtg ctcagcccca acaacttttt tctttctgaa	1260
gtgaaagggc ttaaaatttc ttggaatatag ttttacaaaa atggatttaa aaaatcctac	1320
cgatcaagat gagttcagct agaagtcata ccacctcag gaatcagcta aagcaaaaag	1380
aacttttacc tcggcatcca gccaaccccc taaagactga caatatcctt cgagctcctt	1440
tgaaagcacc ctaaacagcc atttccattt taatagttgg atgcggattg tacccttcaa	1500
tctgaaagtc ttcagctttg aagtcataaa ttttctcaac ttttcgaaga atcctgagct	1560
ttgggaaagg tctgggttct cgctgaagct aaaaacaaaa taaggccatt attttgccat	1620
aattgtacga cctgttgtaa ttgctcctca tgtccgtgaa acaagtacac aggatgtgat	1680
caacaaagtt ctattttaca ggagtatgat cctgtcgata ccttgccgta ggttatgtaa	1740
catgattgga gcgcaaccag ctgttctctt gcacagatcg agagtgaggg gtattttgtg	1800
acattacaca gcacaggag cctggtgcct catcaggtaa ctgaagagct cagtgtggta	1860
ggctggcaag tcaccttcc cgagacagcc caccttcagg cccgtgatgt gcgcaatcat	1920
gtacgtgagc agggcgtagc tggcgatgtt gaaaggcaca ccgaggccca tgtctccga	1980
tctctggtac agctggcagg acagctcact gttcaccaca tagaactggc agagggcatg	2040
gcattggaggc agcgccatca gaggaagatc tcttggttc caagcgaca tgatgattct	2100
tctgtcgta gggttgggtt tgatggtgtc aatcactctt tgcagttggt caactcctg	2160
tctgaataa tctgattcca tatctctgta ttctgcccc aaatgcctcc actggaagcc	2220
ataaactggg cccaagtccc cttcttctct ggtggagaat cccaggctgt ccaaaaagtc	2280
tcgggatcca ttggcatccc agattttcac tcccttgga gacagctctt tagcatttgt	2340
ggatcccttg ataaaccaca gcaactcctc caaaacaccc ttccagaaca cacgtttggt	2400
tgtcagcaga gggaattcat ctctcaggct gtagcgcgcc tgcattgccga ataccgacag	2460
ggtgccggtg cccgtgcggt cgtccttctt gacgccgag cggaggatgt gttggatctg	2520
ccccaggtag tgcagctccc cgtgcggcgg acgcggctcg gcgtcccgt cctgtgcggc	2580
ggggggcaag ggccggcgcg gcagctccga gccggccaca ggcatggcg gcggggcggg	2640
ggacggaggc aggcgaagtg gcgcggcggg acggaggcag gccaaagtggc gcggcgggac	2700
ggaggcaggc caagtggcg ggtggcagga ccctcgncg attaacggag cctcgtaacta	2760
gcca	2764

<210> 61
<211> 3436
<212> DNA
<213> Homo sapien

<400> 61
gccatggtaa gcgcggacgc catggttaagc gcggacgcc tggatatggtta agcgcggacg 60
ccatggtaag cgcggaacgcc atggtaagcg cggacgccat ggtaagcgcg gacgccatgg 120
taagcgcgga cgccatggta agcgcggacg ccatggtaag cgcggaacgcc atggtaagcg 180
cggacgccat ggtaagcgcg gacgccatgc acacggaccc tgactactcg gctgcctatg 240
tcgtcataga aactgatgca gaagatggaa tcaaggggtg tggaattacc ttcactctgg 300
gaaaaggcac tgaagttggt gtctgtgctg tgaatgccct cgcccacccat gtgtctcaaca 360
aggacctcaa ggacattggt ggtgacttca gaggcttcta taggcagctc acaagtgatg 420
ggcagctcag atggattggt ccagaaaagg gcgtggtgca cctggcgaca gcggccgtcc 480
taaacgcggt gtgggacttg tgggccaagc aggagggaaa gcctgtctgg aagttacttg 540
tggacatgga tcccaggatg ctggtatcct gcatagattt caggtacatc actgatgtcc 600
tgactgagga ggatgcccta gaaatactgc agaaagggtca aattggtaaa aaagaaagag 660
agaagcaaat gctggcacia ggataccctg cttacacgac atcgtgcgcc tggctgggggt 720
actcagatga cacgttgaag cagctctgtg ccagggcgct gaaggatggc tggaccagggt 780
ttaaagtaaa ggtgggtgct gatctccagg atgacatgcg aagatgccaa atcatccgag 840
acatgattgg accggaaaag actttgatga tggatgccaa ccagcgctgg gatgtgcctg 900
aggcgggtgga gtggatgtcc aagctggcca agttcaagcc attgtggatt gaggagccaa 960
cctcccctga tgacattctg gggcacgcca ccatttccaa ggactgggtc ccattaggaa 1020
ttggcattgc cacaggagaa cagtgcaca atagagtgat atttaagcaa ctctacagg 1080
cgaaggccct gcagttcctc cagattgaca gttgcagact gggcagtgtc aatgagaacc 1140
tctcagtatt gctgatggcc aaaaagtttg aaattcctgt ttgccccat gctggtggag 1200
ttggcctctg tgaactgggtg cagcacctga ttatatttga ctacatatca gtttctgcaa 1260
gccttgaaaa tagggtgtgt gagtatgtg accacctgca tgagcatttc aagtatcccc 1320
tgatgatcca gcgggcttcc tacatgcctc ccaaggatcc cggctactca acagaaatga 1380
aggaggaatc tgtaaagaaa caccagtatc cagatgggtga agtttggag aaactccttc 1440
ctgctcaaga aaattaagtg ctacgcccc acaacttttt tctttctgaa gtgaaagggc 1500
ttaaattttc ttggaaatag ttttcaaaaa atggatttaa aaaatcctac cgatcaagat 1560
gagttcagct agaagtcata ccacctcag gaatcagcta aagcaaaaag aacttttacc 1620
tcggcatcca gcccaacccc taaagactga caatatacctt cgagctcctt tgaaagcacc 1680

ctaaacagcc atttccattt taatagttgg atgcggattg tacccttcaa tctgaaagtc	1740
ttcagctttg aagtcacaa ttttctcaac ttttcgaaga atcctgagct ttgggaaagg	1800
tctgggttct cgctgaagct aaaaacaaaa taaggccatt attttgccat aattgtacga	1860
cctgttgtaa ttgtcctca tgtccgtgaa acaagtacac aggatgtgat caacaaagtt	1920
ctattttaca ggagtatgat cctgtcgata ccttgccgta ggttatgtaa catgattgga	1980
gcgcaaccag ctgttctctt gcacagatcg agagtgaggg gtattttgtg acattacaca	2040
gcacaggag cctggtgcct catcaggtaa ctgaagagct cagtgtggtg ggctggcaag	2100
tcacccttcc cgagacagcc caccttcagg cccgtgatgt gcgcaatcat gtacgtgagc	2160
agggcgtagc tggcgatggt gaaaggcaca ccgaggccca tgtctcccga tctctggtac	2220
agctggcagg acagctcact gttcaccaca tagaactggc agagggcatg gcacggaggc	2280
agcgccatca gaggaagata ggaaccagca gaggaagata aggagggatg gtggtttgaa	2340
agaccacagc taaaggcaaa gtaaaacagg agagaaacag aagccaactc atatggtgga	2400
gaccaggaga gagagccact gggctgcagt gatgtccata acagcctctg cagcgatggc	2460
acggagctga gggagactat ccacgggtgc aaggtttctg cagggtgtcca ttacggctg	2520
aagcaatgct cttccatcag agctgaaggg atctgggcta cctcgtggca ccagattaca	2580
aatacagcag gaataattct gtttgccaca ggaaactggg gcttctggta caccctccta	2640
tattaaaagt ctctattaca tggccaggca cagtgggtca tgcctgaaat cccaacactg	2700
ggacgccagg gtgggcagat cacttgaggc caacttgagt tcgagactag cctggccaac	2760
atagtgaac cccgtctctg ctaaaaatac aaaaattcgc gcgcctgtgg tcccagctct	2820
ggggctgagg aatgggaact cacttgaacc tgggagatgg aggttgccgt gagctggagt	2880
gcaatggcgt gatcatgact cactacagcc tcaaattcct gggttcaaata gatcctctcg	2940
ccttggcctc tggagtgtgt ggaattacag gtgtgagcca ccacacctgg ctctacaccc	3000
tggtagagtt gaattttaca atataattgt taaaaatgtt agtgaggtgg cagcgagctg	3060
gaacttcac tgcctccatt taaaatgtta acatcttaca tgtagtttct aggcaattta	3120
aaaccctttt tgggtgaaga atgagtataa atacacaaat taagatcatc tgggtgggag	3180
gctgaggcag gcagatcatc tgaggtcagg agttcaagac cagcctgacc aacatggtga	3240
aacccctct ctactaaaaa taaaaatta gccgggcatg gtggcacatg cctgtaatcc	3300
cagctacttg ggaggctgag gcaggagaat cactcgaacc caggaggcag aggttacagt	3360
gagccaagat cagccgttg cattccagcc tgggcaacag agcaaaactc tgtctcaaaa	3420
aaaaaaaaa aaaaaa	3436

<210> 62
 <211> 2928
 <212> DNA
 <213> Homo sapien

<400> 62
 gccatggtaa gcgcggacgc catggtaagc gcggacgcca tggatatgta agcgcggacg 60
 ccatggtaag cgcggaacgc atggtaagcg cggacgcat ggtaagcgcg gacgccatgg 120
 taagcgcgga cgccatggta agcgcggacg ccatggtaag cgcggaacgc atggtaagcg 180
 cggacgcat ggtaagcgcg gacgccatgc acacggaccc tgactactcg gctgcctatg 240
 tcgtcataga aactgatgca gaagatggaa tcaaggggtg tgggaattacc ttcactctgg 300
 gaaaaggcac tgaagttgtt gtctgtgctg tgaatgccct cgcccacccat gtgctcaaca 360
 aggacctcaa ggacattgtt ggtgacttca gaggcttcta taggcagctc acaagtgatg 420
 ggcagctcag atggattggc ccagaaaagg gcgtggtgca cctggcgaca gcggccgtcc 480
 taaacgcggt gtgggacttg tgggccaagc aggagggaaa gcctgtctgg aagtacttg 540
 tggacatgga tcccaggatg ctggtatcct gcatagattt caggtacatc actgatgtcc 600
 tgactgagga ggatgcccta gaaatactgc agaaagggtca aattggtaaa aaagaaagag 660
 agaagcaaat gctggcaca gataccctg cttacacgac atcgtgcgcc tggctggggt 720
 actcagatga cacgttgaag cagctctgtg ccagggcgt gaaggatggc tggaccaggc 780
 ttaaagtaaa ggtgggtgct gatctccagg atgacatgca aagatgcaa atcatccgag 840
 acatgattgg accggaaaag actttgatga tggatgcaa ccagcgtgga gatgtgcctg 900
 aggcggtgga gtggatgtcc aagctggcca agttcaagcc attgtggatt gaggagcaa 960
 cctcccctga tgacattctg gggcacgcca ccatttccaa ggcactgggc ccattaggaa 1020
 ttggcattgc cacaggagaa cagtgcaca atagagtgat atttaagcaa ctccacagg 1080
 cgaaggccct gcagttctc cagattgaca gttgcagact gggcagtgtc aatgagaacc 1140
 tctcagatg gctgatggc aaaaagtgtg aaattcctgt ttgccccat gctggtggag 1200
 ttggcctctg tgaactggtg cagcacctga ttatatttga ctacatatca gtttctgcaa 1260
 gccttgaaaa tagggtgtgt gagtatgtt accacctgca tgagcatttc aagtatcccg 1320
 tgatgatcca gcgggcttc tacatgcctc ccaaggatcc cggctactca acagaaatga 1380
 aggaggaatc tgtaagaaa caccagtatc cagatgggtga agtttgaag aaactccttc 1440
 ctgctcaaga aaattaagtg ctacgcccc acaacttttt tctttctgaa gtgaaagggc 1500
 ttaaaatttc ttggaatag ttttcaaaa atggatttaa aaaatcctac cgatcaagat 1560
 gagttcagct agaagtcata ccacctcag gaatcagcta aagcaaaaag aacttttacc 1620

86

tcggcatcca gcccaacccc taaagactga caatatcctt cgagctcctt tgaaagcacc 1680
 ctaaacagcc atttccattt taatagttgg atgcggattg tacccttcaa tctgaaagtc 1740
 ttcagctttg aagtcaccaa ttttctcaac ttttcgaaga atcctgagct ttgggaaagg 1800
 tctgggttct cgctgaagct aaaaacaaaa taaggccatt attttgccat aattgtacga 1860
 cctgttgtaa ttgctectca tgtccgtgaa acaagtacac aggatgtgat caacaaagtt 1920
 ctattttaca ggagtatgat cctgtcgata ccttgccgta ggttatgtaa catgattgga 1980
 gcgcaaccag ctgttctctt gcacagatcg agagtgaggg gtattttgtg acattacaca 2040
 gcatcaggag cctggtgcct catcaggtaa ctgaagagct cagtgtggta ggctggcaag 2100
 tcacccttcc cgagacagcc caccttcagg cccgtgatgt gcgcaatcat gtacgtgagc 2160
 agggcgtagc tggcgatgtt gaaaggcaca ccgagggcca tgtctcccga tctctggtag 2220
 agctggcagg acagctcact gttcaccaca tagaactggc agagggcatg gcatggaggc 2280
 agcgccatca gaggaagata ggaaccagca gaggaagata aggagggatg gtggtttgaa 2340
 agaccacagc taaaggcaaa gtaaaacagg agagaaacag aagccaactc atatggtgga 2400
 gaccaggaga gagagccact gggctgcagt gatgtccata acagcctctg cagcgatggc 2460
 acggagctga gggagactat ccatcggtgc aagggttctg caggtgtcca tttacggctg 2520
 aagcaatgct cttccatcag agctgaaggg atctgggcta cctcgtggca ccagattaca 2580
 aatacagcag gaataattct gtttgccaca ggaaactggg gcttctggta caccctccta 2640
 tattaaaagt ctctattaca tggccaggca cagtggctca tgcctgaaat cccaacactg 2700
 ggacgccagg gtgggcagat cacttgaggc caacttgagt tcgagactag cctggccaac 2760
 atagtgaac cccgtctctg ctaaaaatac aaaaattcgc gcgcctgtgg tcccagctct 2820
 ggggctgagg aatgggaact cacttgaacc tgggagatgg aggttgccgt gagcggagat 2880
 cacaccactg cactccagcc tgggcaacag ggtgagactg tgtctcaa 2928

<210> 63
 <211> 3100
 <212> DNA
 <213> Homo sapien

<400> 63
 gccatggtaa gcgcggacgc catggtaagc gcggacgcca tggataggta agcgcggacg 60
 ccatggtaag gcgcggacgc atggtaagcg cggacgcat ggtaagcgcg gacgcatgg 120
 taagcgcgga cgccatggta agcgcggacg ccatggtaag gcgcggacgc atggtaagcg 180
 cggacgcat ggtaagcgcg gacgcatgc acacggaccc tgactactcg gctgcctatg 240
 tcgtcataga aactgatgca gaagatggaa tcaaggggtg tggaattacc ttcactctgg 300

gaaaaggcac tgaagttgtt gtctgtgctg tgaatgccct cccccaccat gtgctcaaca	360
aggacctcaa ggacattgtt ggtgacttca gaggcttcta taggcagctc acaagtgatg	420
ggcagctcag atggattggt ccagaaaagg gcgtggtgca cctggcgaca gcggccgtcc	480
taaacgcggt gtgggacttg tgggccaagc aggagggaaa gcctgtctgg aagtacttg	540
tggacatgga tcccaggatg ctggtatcct gcatagattt caggtacatc actgatgtcc	600
tgactgagga ggatgcccta gaaatactgc agaaaggcca aattggtaaa aaagaaagag	660
agaagcaaat gctggcacia ggataccctg cttacacgac atcgtgcgcc tggctgggggt	720
actcagatga cacgttgaag cagctctgtg cccaggcgct gaaggatggc tggaccagggt	780
ttaaagtaaa ggtgggtgct gatctccagg atgacatgcg aagatgcaa atcatccgag	840
acatgattgg accgaaaaag actttgatga tggatgcaa ccagcgctgg gatgtgcctg	900
aggcgggtgga gtggatgtcc aagctggcca agttcaagcc attgtggatt gaggagcaa	960
cctccctga tgacattctg gggcacgcca ccatttccaa ggcaactggc ccattaggaa	1020
ttggcattgc cacaggagaa cagtgccaca atagagtgat atttaagcaa ctccctacagg	1080
cgaaggccct gcagttcctc cagattgaca gttgcagact gggcagtgtc aatgagaacc	1140
tctcagtatt gctgatggcc aaaaagtttg aaattcctgt ttgccccat gctggtggag	1200
ttggcctctg tgaactggtg cagcacctga ttatatattga ctacatatca gtttctgcaa	1260
gccttgaaaa tagggtgtgt gagtatgtt accacctgca tgagcatttc aagtatcccg	1320
tgatgatcca gcgggcttcc tacatgcctc ccaaggatcc cggtactca acagaaatga	1380
aggaggaatc tgtaaagaaa caccagtatc cagatggtga agtttggaag aaactccttc	1440
ctgctcaaga aaattaagtg ctacgcccc acaacttttt tctttctgaa gtgaaagggc	1500
ttaaaatttc ttgaaatag ttttcaaaa atggatttaa aaaatcctac cgatcaagat	1560
gagttcagct agaagtcata ccacctcag gaatcagcta aagcaaaaag aacttttacc	1620
tcggcatcca gccaacccc taaagactga caatatacct cgagctcctt tgaaagcacc	1680
ctaaacagcc atttccattt taatagtgtg atgcggattg tacccttcaa tctgaaagtc	1740
ttcagctttg aagtcataaa ttttctcaac ttttcaaga atcctgagct ttgggaaagg	1800
tctgggttct cgctgaagct aaaaacaaaa taaggccatt attttgccat aattgtacga	1860
cctgttgtaa ttgctctca tgtccgtgaa acaagtacac aggatgtgat caacaaagtt	1920
ctattttaca ggagtatgat cctgtcgata ccttgccgta ggttatgtaa catgattgga	1980
gcgcaaccag ctgttctctt gcacagatcg agagtgaggg gtattttgtg acattacaca	2040
gcatcaggag cctggtgcct catcaggtaa ctgaagagct cagtgtggtg ggctggcaag	2100
tcacccttcc cgagacagcc caccttcagg cccgtgatgt gcgcaatcat gtacgtgagc	2160

```

agggcgtagc tggcgatggt gaaaggcaca ccgaggccca tgtctcccga tctctggtag 2220
agctggcagg acagctcact gttcaccaca tagaactggc agagggcatg gcatggaggc 2280
agcgccatca gaggaagatc tcttggattc caagcgcaca tgatgattct tctgtcgtca 2340
gggttggttt tgatggtgtc aatcactctt tgcagttggt caactccctg tcctgaataa 2400
tctgattcca tatctctgta ttctgcccc aatgcctcc actggaagcc ataaactggg 2460
cccaagtccc cttcttctct ggtggagaat cccaggctgt ccaaaaagtc tcgggatcca 2520
ttggcatccc agattttcac tcccttgga gacagctctt tagcatttgt ggatcccttg 2580
ataaaccaca gcaactctc caaaacaccc ttccagaaca cacgtttggt tgtcagcaga 2640
gggaattcat ctctcaggct gtagcgcgcc tgcattgccga ataccgacag ggtgccggtg 2700
cccgtgcggt cgtccttcct gacgccgcag cggaggatgt gttggatctg ccccaggatt 2760
tccagggtccc agatgaagag ataattctac ttactggata taggatgcat tagatcttct 2820
taccttaaaa aaaaaaaaaa ggcagcaatg atcaaaatac taataaatta ctacagact 2880
cagtgtatct tttcttgagg taaaagtcca ggatgggtaa tagaatacct gctgttggtc 2940
tttggaaaaa ttggtactgt atgtagcaaa ataattgtgaa acccatatgc atggatatct 3000
ttaacaatct gaagaaatcg tcacagcttt cctgggttgt tgagcctcta agatggtctt 3060
ttcctctgat gtgataataa agtggtttatt ctgaactcta 3100

```

<210> 64
 <211> 463
 <212> DNA
 <213> Homo sapien

```

<400> 64
tggatcgtag tcgcggcgag gtacaggatt agactggggt tatatttaac tcttgcctca 60
taggtgtact acacgcgcct gggcaacgac ttccacacga acaagcgct gtgcgaggag 120
atcgccatta tccccagcaa aaagctccgc aacaagatag caggttatgt cacgcattctg 180
atgaagcgaa ttcagagagg ccagtaaga ggtatctcca tcaagctgca ggaggaggag 240
agagaaagga gagacaatta tgttcctgag gtctcagcct tggatcagga gattattgaa 300
gtagatcctg acactaagga aatgctgaag cttttggact tcggcagtct gtccaacctt 360
caggctactc agcctacagt tgggatgaat ttcaaaacgc ctgggggacc tgtttgaatt 420
ttttctgtag tgctgtatta ttttcaataa atctgggaca aca 463

```

<210> 65
 <211> 5226
 <212> DNA
 <213> Homo sapien

<400> 65
 gcggtggcgg tggcggccgc gtcccggcgg cagtcgtgct acctgtgtga cctgccccgc 60
 atgccctggg ccatgatctg ggacttcacc gaacccgtct gccgcggctg cgtcaactac 120
 gagggcgccg accgcgtcga gtctgtcatc gagacggcgc ggcagctcaa gggggcgcac 180
 ggctgcttcc cggagggtcg ctccccaccc ggcgccgcgg cctcggccgc cgccaagccg 240
 ccgcgcgtct ccgccaagga catccttttg cagcagcagc agcagcttgg ccacggcggc 300
 cccgaggcgg ccccgcgcg cccgcaggcc ttggagcgt acccgttggc ggccgcggcc 360
 gagaggcccc cgcgcctcgg ctctgacttc ggagcagcc gcccggcagc gagcctggcc 420
 cagccgcga cgccgcagcc gccgcccgtg aacggcatcc tggtgccaa cggttctcc 480
 aagctagagg agccgcccga gctgaatcgc cagagcccga acccgcgcg cgccacgcg 540
 gtgccgcca ccctggtgcc gctcatgaac ggctcggcca cgccgctgcc caccgcgtc 600
 ggctcggcg gccgcgtgc cgcctcctta gccgcggtgt ccggaaccgc ggccgccagc 660
 ctgggctccg cgcagccac cgatctgggc gccacaagc ggccggcatc cgtgtcgagc 720
 agcgtgccg tggagcacga gcagcgtgag gcggcagcca aggagaaaca accgccgccg 780
 cctgcgcacc ggggcccggc cgacagcctg tccaccgcgg ccggggccgc cgagctgagc 840
 gcggaagggt cgggcaagag ccgcgggtct ggagagcagg actgggtcaa caggcccaag 900
 accgtgcgcg acacgctgct ggcgctgcac cagcacggcc actcggggcc cttcgagagc 960
 aagtttaaga aggagccggc cctgactgca ggcaggttgt tgggtttcga ggccaacggg 1020
 gccaacgggt ctaaagcagg taggggcggc tgtgaagtga gggggtctag gggagaaaag 1080
 gggacggaga gcagaggaag ggtggttctt tggattcacc attttaccac agcccagaaa 1140
 caacaaacac ccacttcct gatctgcctg aggcggaacc agtgcttggg ggcaacgtgt 1200
 tcatgtgctg aagcagcata acagagatga gtcagactgg gctgatacgc tctgacacgg 1260
 gggtttcctt tcccagcaca ttcttgatg ggagcatgag ggcaccagtc accttttaac 1320
 ctattggggg acattagcag tcacatggtg agtgcaaagc agggctacttt tgtgcatgtg 1380
 taaaaacagg cagttacaag cgtgtcattt tcagtggctc cattttaaat cagtctgctg 1440
 cctcagaatc ccgtacgcct gaaggtttta agttgcatgt gcacctgaaa ctcttatatg 1500
 agtattttct gtctgtgctt ttagagagga ggaattctgt aacgactttt gtttcggggt 1560
 aggaagagaa tgatctcttt cagtgcaccg ccacttatgt tacctttttc cttttatttc 1620
 tttgtgttcc cagttgcaag aacagcaagg aaaaggaagc cctctccaga accagaaggt 1680
 gaagtcgggc cccctaagat caacggagag gcccgccgt ggctgtccac atccacagag 1740
 gggctcaaga tccccatgac tcctacatcc tcttttgtgt ctccgccacc acccactgcc 1800

tcacctcatt ccaaccggac cacaccgcct gaagcggccc agaatggcca gtcccccatg	1860
gcagccctga tcttagtagc agacaatgca gggggcagtc atgcctcaaa agatgccaac	1920
caggttcact ccactaccag gaggaatagc aacagtccgc cctctccgtc ctctatgaac	1980
caaagaaggc tgggccccag agaggtgggg ggccaggagg caggcaacac aggaggactg	2040
gagccagtgc accctgccag cctcccggac tcctctctgg caaccagtgc cccgtgtg	2100
tgcacctct gccacgagcg gctggaggac acccattttg tgcagtgcc gtccgtccct	2160
tcgcacaagt tctgcttccc ttgctccaga caaagcatca aacagcagg agctagtga	2220
gaggtctatt gtcccagtgg ggaaaaatgc cctcttgg gctccaatgt cccctggg	2280
tttatgcaag gggaaattgc aaccatcctt gctggagatg tgaaagtga aaaagagaga	2340
gactcgtgac ttttccggtt tcagaaaaac ccaatgatta cccttaatta aaactgctt	2400
aattgtatat atatctccat atatatatat atccaagaca agggaaatgt agacttcata	2460
aacatggctg tataattttg atttttttg aatacattgt gtttctatat ttttttgac	2520
gacaaaagg atgtacttat aaagacattt ttttctttg ttaacgttat tagcatatct	2580
ttgtgcttta ttatcctggt gacagttacc gttctatgta ggctgtgact tgcgtgctt	2640
ttttagagca cttggcaaat cagaaatgct tctagctgta tttgtatgca cttatttta	2700
aaagaaaaaa aaagccaaat acattttctg acattgtaag attgccttac tgtctgcat	2760
tccttattgc tggcccttt ctgagccgg aggccaagtg gtggagaagg aaaggaaatg	2820
atcgaacggg catgttgtca agtgggcatg ccactgggaa ataccaccag tttaccctga	2880
aacattgtcc tcagaggagt aggaaagtgg attttgaatc tctattttgc tcaaaagttc	2940
agttcctgag atactgatga ctgagagtgc tgctgggaaa ttttcaggat tgtgtggtc	3000
tttggggttt tttgttttt tttttttaag acaaagttga ccgctgttca ctgtccacgt	3060
gatcagttgt aagattacaa tgctgcatgc tagttggtta cataagatac aattccagt	3120
atggaaggcg gttataatgg atggtggtgt gtacaagatg gcaactgccat ctttgagcag	3180
agcccagctc tgcagcgcca cttcatctt ttaaacaccc tagaggtctg tttgttggtg	3240
ctgttgctct ttattttgaa agagttgcaa gagaagttac agtccagggtg aacttgaga	3300
ttgtgggatt gggtttgtt ctgtttgtt ttgtttatca tttacctgta gtgctattgc	3360
tggtgatact atcacctata ccctgttct agtgagtgt gaatacagta tggtaaatg	3420
acagtaacag ccgctgggtg ctgccaggac tgccttggg catatcagt acagcccaa	3480
tgtgggtgga ggaaacctgt aatttcctt ttaacatgt tttgaaatac caagtgaata	3540
atactgttct ggaaaaaat gataaactag tggaaattaa agaaattaag gggtttatat	3600

91

```

aatagacagg cccacacctt caaaatattt ttagaagctt ttttgtaaac taatttcttt 3660
tgatcactat tttgcatcag taaaatgatt tttttaaaac caataaatca tcaattatta 3720
gaaatagttg tctcacagtg atactgggtt ttcttttggt ctgttatgat ttaacattga 3780
caggaacact attttaaatc cttacgttca ggtgtttgta acttggcctt ataattaggg 3840
tgaattatgg cttcaaggtc tacaatttat gtgtatgggt cacagcctag cttctattta 3900
catttgaaaa tacagatttt taccaacttt ggattctttt ttagttatat gtttgtcttt 3960
ccttttttaa ttgttcaaaa ctatttttta atggtcaagt tactaacact tgaaaatcag 4020
atactgcacc aaatacagtg tttttccgta gtgtttttaa tgagtgcacc tattactact 4080
gtgcgagaat tcatgtttta ccagtcattg ttatattaca aacagacttg catgattaac 4140
cagttgttac acttactttt tcaagttgga gtatatatga ctcagtgcag actggtctct 4200
cttatgtgaa tgcacacatg cagaaatgca gagtcaattt tacatgccca taaagacatt 4260
tgtaaagaat tcagctctta tggctctgtg tataaatgtg tatctaggca ctttattata 4320
actggaattt gacctcacag atgttacaac ttgatcagtc gtttgacctt atttgtggta 4380
gctatctgta tgttttgcaa tcttaataca gacatgcttt ccaaaaagat taatacagaa 4440
ccatcctgcc gttttggata agtctatcca gctgtggaaa gggcaacctg tggtttctct 4500
gtactggtgt ttaatggggg aagaatatga acagctttaa agagctgtgt attgtgggta 4560
ctactattaa aaaataagat ctgcacgagt ctgactggcc tttgggtggc ctttgtggac 4620
ggctcgtagc tggaaagtgt tgatctgggt tttctggcat tcttttaagt taaaaagtta 4680
acatcgggac atgggtttga tcttttggtg tacctgatga cagtgcagag attctccaca 4740
gctggataaa aatgtcacia agctacttac tgtacatggg cagtatcaga tttcaaattc 4800
taatatttca gctgtgcttt taatactcaa aatattaggg gatgggggtg tgaagctttc 4860
ccttttttgc ttttaacaat ttatagaatt taacagatgt actgtctttc atgtggcctc 4920
acatttaaag ttatgagaac atacacatgg tttaacaact ttactatata cctttccttg 4980
gccaccaagt attttaaaag tgtgccacct tttaaccttt acttttttta agtgaaggt 5040
gatacttttt ctatatatga tgaaactcat gtcaactgaa gtgagtgtaa tctcagatat 5100
caacattatt atattttaaa atcacgctat ggaaatatca cctgaattct gtcatttgtc 5160
agatttacag tacctttttt tctttaactt ttagcattaa ataaaaataa aattgggagc 5220
actgaa 5226

```

<210> 66
<211> 5045
<212> DNA
<213> Homo sapien

<400> 66
 gcggtggcgg tggcggccgc gtcccggcgg cagtctgtgt acctgtgtga cctgccccgc 60
 atgccctggg ccatgatctg ggacttcacc gaaccctgtct gccgcggctg cgtcaactac 120
 gagggcgccg accgcgtcga gtctgtcatc gagacggcgc ggcagctcaa gcggggcgac 180
 ggctgcttcc cggagggctg ctccccaccc ggcgcgcgg cctcggccgc cgccaagccg 240
 ccgcccgtct ccgccaagga catccttttg cagcagcagc agcagcttg ccacggcggc 300
 cccgagggcg ccccgcgcg gccgcaggcc ttggagcgt acccgttggc ggccgcggcc 360
 gagaggcccc cgcgccctcg ctctgacttc ggagcagcc gccggcgagc gagcctggcc 420
 cagccgccga cgccgcagcc gccgcccgtg aacggcatcc tggtgcccaa cggcttctcc 480
 aagctagagg agccgcccga gctgaatcgc cagagcccga acccgcggcg cggccacgcg 540
 gtgccgcccc ccctgggtgcc gctcatgaac ggctcggcca cgccgctgcc caccgcgctc 600
 ggccctcggcg gccgcgctgc cgcctcctta gccgcgggtgt ccggaaccgc ggccgccagc 660
 ctgggctccg cgcagcccac cgatctgggc gccacaagc ggccggcatc cgtgtcgagc 720
 agcgtgccc tggagcacga gcagcgtgag gcggcagcca aggagaaaca accgccgccc 780
 cctgcgcacc gggggcccgc cgacagcctg tccaccgcgg ccggggccgc cgagctgagc 840
 gcggaagggt cgggcaagag ccgcgggtct ggagagcagg actgggtcaa caggcccaag 900
 accgtgcgcg acacgctgct ggcgctgcac cagcacggcc actcggggcc cttcgagagc 960
 aagtttaaga aggagccggc cctgactgca ggaggttgt tgggtttcga ggccaacggg 1020
 gccaacgggt ctaaagcagg tagggcgggc tgtgaagtga ggggggtctag gggagaaaag 1080
 gggacggaga gcagaggaag ggtggttctt tggattcacc attttaccac agcccagaaa 1140
 caacaaacac ccacttctct gatctgcctg aggcggaacc agtgcttggg ggcaacgtgt 1200
 tcatgtgctg aagcagcata acagagatga gtcagactgg gctgatacgc tctgacacgg 1260
 gggttttctt tcccagcaca ttcttggtg ggagcatgag ggcaccagtc accttttaac 1320
 ctattggggg acattagcag tcacatgttg agtgcaaac aggggtacttt tgtgcatgtg 1380
 taaaaacagg cagttacaag cgtgtcattt tcagtggctc catttttaaat cagtctgctg 1440
 cctcagaatc ccttgcaaga acagcaagga aaaggaagcc ctctccagaa ccagaagggt 1500
 aagtcggggc ccctaagatc aacggagagg cccagccgtg gctgtccaca tccacagagg 1560
 ggctcaagat ccccatgact cctacatcct cttttgtgtc tccgccacca cccactgcct 1620
 cacctcattc caaccggacc acaccgcctg aagcggccca gaatggccag tcccccatgg 1680
 cagccctgat cttagtagca gacaatgcag ggggcagtca tgcctcaaaa gatgccaaac 1740
 aggttcactc cactaccagg aggaatagca acagtccgcc ctctccgtcc tctatgaacc 1800

aaagaaggct gggccccaga gaggtggggg gccagggagc aggcaacaca ggaggactgg 1860
agccagtgca ccctgccagc ctcccggact cctctctggc aaccagtgcc ccgctgtgct 1920
gcaccctctg ccacgagcgg ctggaggaca cccattttgt gcagtgcccg tccgtccctt 1980
cgcacaagtt ctgcttcctt tgctccagac aaagcatcaa acagcagga gctagtggag 2040
aggctctattg tcccagtggg gaaaaatgcc ctcttgtggg ctccaatgtc ccctgggcct 2100
ttatgcaagg ggaaattgca accatccttg ctggagatgt gaaagtgaaa aaagagagag 2160
actcgtgact tttccggttt cagaaaaacc caatgattac ccttaattaa aactgcttga 2220
attgtatata tatctccata tatatatata tccaagacaa gggaaatgta gacttcataa 2280
acatggctgt ataattttga ttttttttga atacattgtg tttctatatt ttttttgacg 2340
acaaaaggta tgtacttata aagacatttt tttcttttgt taacgttatt agcatatctt 2400
tgtgctttat tatcctgggtg acagttaccg ttctatgtag gctgtgactt gcgctgcttt 2460
tttagagcac ttggcaaata agaaatgctt ctgctgtat ttgtatgcac ttattttaaa 2520
aagaaaaaaa aagccaaata cattttctga cattgtaaga ttgccttact gtctgtcatt 2580
ccttattgct ggcccccttc tcaggccgga ggccaagtgg tggagaagga aaggaaatga 2640
tcgaacgggc atgttgtcaa gtgggcatgc cactgggaaa taccaccagt ttaccctgaa 2700
acattgtcct cagaggagta ggaaagtgga ttttgaatct ctattttgct caaaagttca 2760
gttcttgaga tactgatgac tgagagtgtg gctgggaaat ttccaggatt gtgtggcttt 2820
ttggggtttt ttgttttttt ttttttaaga caaagttgac cgctgttcac tgtccacgtg 2880
atcagttgta agattacaat gctgcatgct agttggttac ataagataca attccagtga 2940
tggaaggcgg ttataatgga tgggtggtgtg tacaagatgg cactgccatc tttgagcaga 3000
gcccagctct gcagcgccac ttcatctttt taaacaccct agaggtctgt ttgttggtgc 3060
tgttgctcctt tattttgaaa gagttgcaag agaagttaca gtccaggatga acttggagat 3120
tgtgggattg gttttggttc tgttttgttt tgtttatcat ttacctgtag tgctattgct 3180
gttgatacta tcacctatac cctgtttcta gtgagtgtg aatacagtat ggtacaatga 3240
cagtaacagc cgcgtgggtgc tgccaggact gcccttgggc atatcagtga cagcccaaata 3300
gtgggtggag gaaacctgta atttccttct taacatgtgt ttgaaatacc aagtgaataa 3360
tactgttctg gaaaaaatg ataaactagt ggaaattaaa gaaattaagg gttttatata 3420
atagacaggc cccacctctc aaaatatatt tagaagtctt tttgtaaaact aatttccttt 3480
gatcactatt ttgcatcagt aaaatgattt ttttaaaacc aataaatcat caattattag 3540
aaatagttgt ctcacagtga tactggtttt tcttttgtgc tgttatgatt taacattgac 3600

94

```

aggaacacta ttttaaatcc ttacgttcag gtgtttgtaa cttggcctta taattaggct 3660
gaattatggc ttcaaggctt acaatttatg tgtatgggtc acagcctagc ttctatttac 3720
at ttgaaaat acagattttt accaactttg gattcttttt tagttatatg tttgtctttc 3780
cttttttaaat tgttcaaaac ttttttttaa tgggtcaagtt actaacactt gaaaaatcaga 3840
tactgcacca aatacagtgt tttccgtag tgtttttaat gagtgcacct attactactg 3900
tgcgagaatt catgttttac cagtcattgt tatattacaa acagacttgc atgattaacc 3960
agttgttaca cttacttttt caagttggag tatatatgac tcagtgcaga ctggtctctc 4020
ttatgtgaat gcacacatgc agaaatgcag agtcaatttt acatgcccat aaagacattt 4080
gtaaagaatt cagctcttat ggtctgttgt ataaatgtgt atctaggcac tttattataa 4140
ctggaatttg acctcacaga tgttacaact tgatcagtcg tttgacctaa tttgtggtag 4200
ctatctgtat gttttgcaat ctttaatacag acatgctttc caaaaagatt aatacagaac 4260
catcctgccg ttttggataa gtctatccag ctgtggaaaag ggcaacctgt ggtttctctg 4320
tactgggtgt taatggggga agaatatgaa cagctttaaa gagctgtgta ttgtggttac 4380
tactattaaa aaataagatc tgcacgagtc tgactggcct ttgggtggcc tttgtggacg 4440
gctcgtagct ggaaagtgtt gatctgggtt ttctggcatt cttttaagtt aaaaagttaa 4500
catcgggaca tgggtttgat cttttgttgt acctgatgac agtgcagaga ttctccacag 4560
ctggataaaa atgtcacaaa gctacttact gtacatgggc agtatcagat ttcaaactct 4620
aatatttcag ctgtgctttt aatactcaaa atattagggg atgggggtgtt gaagctttcc 4680
cttttttgct ttttaacaatt tatagaattt aacagatgta ctgtctttca tgtggcctca 4740
catttaaaagt tatgagaaca tacacatggg ttacaacttt tactatatac ctttccttgg 4800
ccaccaagta ttttaaaagt gtgccacctt ttaaccttta ctttttttaa gttgaagggtg 4860
atactttttc tatatatgat gaaactcatg tcaactgaag tgagtgtaat ctcatatc 4920
aacattatta ttttttaaaa tcacgctatg gaaatatcac ctgaattctg tcatttgtca 4980
gatttacagt accttttttt ctttaacttt tagcattaaa taaaaataaa attgggagca 5040
ctgaa 5045

```

```

<210> 67
<211> 4956
<212> DNA
<213> Homo sapien

```

```

<400> 67
gcgggtggcgg tggcggccgc gtcccggcgg cagtcgtgct acctgtgtga cctgccccgc 60
atgccctggg ccgatgctg ggacttcacc gaaccgctct gccgcggctg cgtcaactac 120

```

95

gagggcgccg accgcgtcga gtctgtcatc gagacggcgc ggagctcaa gcggcgccac	180
ggctgcttcc cggagggtcg ctccccaccc ggccgcccgc cctcgccgc cgccaagccg	240
ccgccgctct ccgccaagga catccttttg cagcagcagc agcagcttgg ccacggcgcc	300
cccagggcgg ccccgccgc gccgcaggcc ttggagcgt acccgttggc ggccgcccgc	360
gagaggcccc cgcgcctcgg ctctgacttc ggagcagcc gcccgccagc gagcctggcc	420
cagccgcccga cgcgcagcc gccgcccgtg aacggcatcc tggtgccca cggcttctcc	480
aagctagagg agccgcccga gctgaatcg cagagcccga acccgccggc cggccacgcg	540
gtgccgccc cctgggtgcc gctcatgaac ggctcgccca cgcgctgcc caccgcgtc	600
ggcctcgccg gccgcgtgc cgcctcctta gccgcggtgt ccggaaccgc ggccgcccgc	660
ctgggctccg cgcagccac gatctgggc gccacaagc ggccggcatc cgtgtcagc	720
agcgtgccg tggagcacga gcagcgtgag gccgcagcca aggagaaaca accgcccgcg	780
cctgcgcacc ggggcccggc cgacagcctg tccaccgcgc ccggggccgc cgagctgagc	840
gcggaagggt cgggcaagag ccgcccgtct ggagagcagg actgggtcaa caggcccaag	900
accgtgcgcg acacgctgct ggcgtgcac cagcacggcc actcggggcc ctctgagagc	960
aagttaaga aggagccggc cctgactgca ggaggttgt tgggtttcga ggccaacggg	1020
gccaacgggt ctaaagcagg tagggcgcc tgtgaagtga ggggtctag gggagaaaag	1080
gggacggaga gcagaggaag ggtggttctt tggattcacc attttaccac agcccagaaa	1140
caacaaacac ccacttcct gatctgcctg aggcggaacc agtgcttggg ggcaacgtgt	1200
tcatgtgctg aagcagcata acagagatga gtcagactgg gctgatacgc tctgacacgg	1260
ggttttcctt tcccagcaca ttcttgatg ggagcatgag ggcaccagtc accttttaac	1320
ctattggggg acattagcag tcacatgttg agtgcaaacg aggttgcaag aacagcaagg	1380
aaaaggaagc cctctccaga accagaagg gaagtcgggc ccctaagat caacggagag	1440
gcccagccgt ggctgtccac atccacagag gggctcaaga tcccacatgac tcctacatcc	1500
tcttttgtgt ctccgccacc acccactgcc tcacctcatt ccaaccggac cacaccgcct	1560
gaagcggccc agaatggcca gtcccccag gcagccctga tcttagtagc agacaatgca	1620
gggggcagtc atgcctcaaa agatgccaac caggttcact ccactaccag gaggaatagc	1680
aacagtccgc cctctccgct ctctatgaac caaagaaggc tgggccccag agaggtgggg	1740
ggccagggag caggcaacac agggaggactg gagccagtgc accctgccag cctcccgac	1800
tcctctctgg caaccagtgc cccgctgtgc tgcacctct gccacgagcg gctggaggac	1860
acccattttg tgcagtgcc gtccgtccct tcgcacaagt tctgcttccc ttgctccaga	1920
caaagcatca aacagcaggg agctagtgg gaggtctatt gtcccagtg gaaaaaatgc	1980

cctcttgtgg gctccaatgt cccctgggccc tttatgcaag gggaaattgc aaccatcctt 2040
gctggagatg tgaaagtga aaaagagaga gactcgtgac ttttccggtt tcagaaaaac 2100
ccaatgatta cccttaatta aaactgcttg aattgtatat atatctccat atatatatat 2160
atccaagaca agggaaatgt agacttcata aacatggctg tataattttg attttttttg 2220
aatacattgt gtttctatat tttttttgac gacaaaagg atgtacttat aaagacattt 2280
ttttcttttg ttaacgttat tagcatatct ttgtgcttta ttatcctggg gacagttacc 2340
gttctatgta ggctgtgact tgcgctgctt ttttagagca cttggcaaat cagaaatgct 2400
tctagctgta tttgtatgca cttattttta aaagaaaaaa aaagccaaat acattttctg 2460
acattgtaag attgccttac tgtctgtcat tccttattgc tggccccctt ctcaggccgg 2520
aggccaagt gtggagaagg aaaggaaatg atcgaacggg catgttgtca agtgggcatg 2580
ccactgggaa ataccaccag tttaccctga aacattgtcc tcagaggagt aggaaagtgg 2640
attttgaatc tctattttgc tcaaaagttc agttcctgag atactgatga ctgagagtgc 2700
tgctgggaaa ttttcaggat tgtgtggtct tttggggttt tttgtttttt tttttttaag 2760
acaaagtga ccgctgttca ctgtccacgt gatcagttgt aagattacaa tgtgcatgc 2820
tagttgggta cataagatac aattccagt atggaaggcg gttataatgg atggtgggtg 2880
gtacaagatg gcactgccat ctttgagcag agcccagctc tgcagcgcca cttcatcttt 2940
ttaaacaccc tagaggctg tttgtgttg ctgttgcct ttattttgaa agagttgcaa 3000
gagaagttac agtccagggt aacttgaga ttgtgggatt ggttttgttt ctgttttgtt 3060
ttgtttatca tttacctgta gtgctattgc tgttgatact atcacctata ccctgtttct 3120
agtgagtgt gaatacagta tggtaaatg acagtaacag ccgctgggtg ctgccaggac 3180
tgcccttggg catatcagt acagcccaaa tgtgggtgga ggaaacctgt aatttccttc 3240
ttaacatgtg tttgaaatac caagtgaata atactgttct ggaaaaaat gataaactag 3300
tggaattaa agaaattaag ggttttatat aatagacagg cccacctct caaaatattt 3360
ttagaagtct tttgtaaac taatttcttt tgatcactat tttgcatcag taaaatgatt 3420
tttttaaac caataaatca tcaattatta gaaatagttg tctcacagt atactggttt 3480
ttcttttgtg ctgttatgat ttaacattga caggaacact attttaaat cttacgttca 3540
gggttttgtg acttggcctt ataattaggc tgaattatgg cttcaaggtc tacaatttat 3600
gtgtatgggt cacagcctag cttctattta catttgaaaa tacagatttt taccaacttt 3660
ggattctttt ttagttatat gtttgcctt cctttttaaa ttgttcaaaa ctatttttta 3720
atggtcaagt tactaacact tgaaaatcag atactgcacc aaatacagt ttttccgta 3780

97

gtgttttttaa tgagtgcacc tattactact gtgcgagaat tcatgtttta ccagtcattg 3840
 ttatattaca aacagacttg catgattaac cagttgttac acttactttt tcaagttgga 3900
 gtatatatga ctcaagtgcag actggtctct cttatgtgaa tgcacacatg cagaaatgca 3960
 gagtcaattt tacatgcccc taaagacatt tgtaaagaat tcagctctta tgggtctgtg 4020
 tataaatgtg tatctaggca ctttattata actggaattt gacctcacag atgttacaac 4080
 ttgatcagtc gtttgaccta atttgtggtg gctatctgta tgttttgcaa tcttaataca 4140
 gacatgcttt ccaaaaagat taatacagaa ccatcctgcc gttttggata agtctatcca 4200
 gctgtggaaa gggcaacctg tggtttctct gtactggtgt ttaatggggg aagaatatga 4260
 acagctttta agagctgtgt attgtggtta ctactattaa aaaataagat ctgcacgagt 4320
 ctgactggcc tttgggtggc ctttgtggac ggctcgtagc tggaaagtgt tgatctgggt 4380
 tttctggcat tcttttaagt taaaaagta acatcgggac atgggtttga tcttttgttg 4440
 tacctgatga cagtgcagag attctccaca gctggataaa aatgtcacaag agctacttac 4500
 tgtacatggg cagtatcaga tttcaaatcc taatatcca gctgtgcttt taatactcaa 4560
 aatattaggg gatggggtgt tgaagctttc ctttttttgc ttttaacaat ttatagaatt 4620
 taacagatgt actgtctttc atgtggcctc acatttaaag ttatgagaac atacacatgg 4680
 tttacaactt ttactatata cctttccttg gccaccaagt attttaaaag tgtgccacct 4740
 ttttaacctt acttttttta agttgaaggt gatacttttt ctatatatga tgaaactcat 4800
 gtcaactgaa gtgagtgtaa tctcagatat caacattatt atattttaaa atcacgctat 4860
 ggaaatatca cctgaattct gtcatttgtc agatttacag tacctttttt tctttaactt 4920
 ttagcattaa ataaaaataa aattgggagc actgaa 4956

<210> 68
 <211> 4442
 <212> DNA
 <213> Homo sapien

<400> 68
 gcggtggcgg tggcggccgc gtcccggcgg cagtcgtgct acctgtgtga cctgccccgc 60
 atgccctggg ccatgatctg ggacttcacc gaaccctct gccgcggctg cgtcaactac 120
 gagggcgccg accgcgtcga gtctgtcatc gagacggcgc ggcagctcaa gcgggcgcac 180
 ggctgcttcc cggaggggtc ctccccaccc ggcgccgcgg cctcggccgc cgccaagccg 240
 ccgccgtct cgcgaagga catccttttg cagcagcagc agcagcttgg ccacggcggc 300
 cccgaggcgg ccccgcgccg gccgcaggcc ttggagcgt acccgttggc ggccgcggcc 360
 gagaggcccc cgcgcctcgg ctctgacttc ggcagcagcc gcccggcagc gagcctggcc 420

cagccgccga cgccgcagcc gccgcccgtg aacggcatcc tggcgcccaa cggcttctcc	480
aagctagagg agccgcccga gctgaatcgc cagagcccga acccgcggcg cggccacgcg	540
gtgccgccca ccctgggtgcc gctcatgaac ggctcggcca cgccgctgcc caccgcgctc	600
ggcctcggcg gccgcgctgc cgcctcctta gccgcggtgt ccggaaccgc ggccgccagc	660
ctgggctccg cgcagcccac cgatctgggc gccacacaagc ggccggcatc cgtgtcgagc	720
agcgtcgccg tggagcacga gcagcgtgag ggcgcagcca aggagaaaca accgccgccc	780
cctgcgcacc ggggcccggc cgacagcctg tccaccgcgg ccggggccgc cgagctgagc	840
gcggaagggtg cgggcaagag ccgcgggtct ggagagcagg actgggtcaa caggcccaag	900
accgtgcgcg acacgctgct ggcgctgcac cagcacggcc actcggggcc cttcgagagc	960
aagttaaaga aggagccggc cctgactgca gttgcaagaa cagcaaggaa aaggaagccc	1020
tctccagaac cagaagggtga agtcggggcc cctaagatca acggagaggc ccagccgtgg	1080
ctgtccacat ccacagaggg gctcaagatc cccatgactc ctacatcctc ttttgtgtct	1140
ccgccaccac ccactgctc acctcattcc aaccggacca caccgcctga agcggcccag	1200
aatggccagt ccccatggc agccctgatc ttagtagcag acaatgcagg gggcagtcac	1260
gcctcaaaag atgccaacca ggttcatact ctctggcaac cagtgcocccg ctgtgctgca	1320
ccctctgcca cgagcggctg gaggacacc attttgtgca gtgcccgtcc gtcccttcgc	1380
acaagtcttg cttcccttgc tccagacaaa gcatcaaaaca gcagggagct agtggagagg	1440
tctattgtcc cagtggggaa aaatgcctc ttgtgggctc caatgtcccc tgggccttta	1500
tgcaagggga aattgcaacc atccttgctg gagatgtgaa agtgaaaaaa gagagagact	1560
cgtgactttt ccggtttcag aaaaacccaa tgattaccct taattaaaac tgcttgaatt	1620
gtatatatat ctccatatat atatatatcc aagacaaggg aaatgtagac ttcataaaca	1680
tggctgtata attttgattt ttttgaata cattgtgttt ctatatTTTT tttgacgaca	1740
aaaggtatgt acttataaag acattttttt cttttgttaa cgttatttagc atatctttgt	1800
gctttattat cctggtgaca gttaccgttc tatgtaggct gtgacttgcg ctgctttttt	1860
agagcacttg gcaaatacaga aatgcttcta gctgtatttg tatgcactta ttttaaaaag	1920
aaaaaaaaag ccaaatacat tttctgacat tgtaagattg ccttactgtc tgtcattcct	1980
tattgctggc ccctttctca ggccggaggc caagtgggtg agaaggaaag gaaatgatcg	2040
aacgggcatg ttgtcaagtg ggcatgccac tgggaaatac caccagttaa ccctgaaaca	2100
ttgtcctcag aggagtagga aagtggattt tgaatctcta ttttgctcaa aagttcagtt	2160
cctgagatac tgatgactga gagtgtgct gggaaaattt caggattgtg tggctctttg	2220
gggttttttg tttttttttt ttttaagacaa agttgaccgc tgttactgt ccacgtgatc	2280

agttgtaaga ttacaatgct gcatgctagt tggttacata agatacaatt ccagtgatgg 2340
aaggcggtta taatggatgg tgggtgtgtac aagatggcac tgccatcttt gagcagagcc 2400
cagctctgca gcgccacttc atctttttaa acaccctaga ggtctgtttg ttgttgctgt 2460
tgtcctttat tttgaaagag ttgcaagaga agttacagtc cagggtgaact tggagattgt 2520
gggattgggt ttgtttctgt tttgttttgt tcatcattta cctgtagtgc tattgctgtt 2580
gatactatca cctataccct gtttctagtg agtgctgaat acagtatggg acaatgacag 2640
taacagccgc gtggtgctgc caggactgcc cttgggcata tcagtgcag cccaaatgtg 2700
gggtggaggaa acctgtaatt tccttcttaa catgtgtttg aaataccaag tgaataatac 2760
tgttctggaa aaaaatgata aactagtggg aattaaagaa attaagggtt ttatataata 2820
gacaggcccc acctctcaa atatttttag aagtcttttt gttaaactaat ttcttttgat 2880
cactattttg catcagtaaa atgatttttt taaaaccaat aaatcatcaa ttattagaaa 2940
tagttgtctc acagtgtac tgggttttct tttgtgctgt tatgatttaa cattgacagg 3000
aacactattt taaatcctta cggtcagggt tttgtaactt ggctttataa ttaggtgaa 3060
ttatggcttc aaggcttaca atttatgtgt atgggtcaca gcctagcttc tatttacatt 3120
tgaaaataca gatttttacc aactttggat tcttttttag ttatatgttt gtctttcctt 3180
tttaaatgt tcaaaactat tttttaatgg tcaagttact aacacttgaa aatcagatac 3240
tgcaccaa at acagtgtttt tccgtagtgt ttttaatgag tgcacctatt actactgtgc 3300
gagaattcat gttttaccag tcattgttat attacaaaca gacttgcag attaaccagt 3360
tgttacactt actttttcaa gttggagtat atatgactca gtgcagactg gtctctctta 3420
tgtgaatgca cacatgcaga aatgcagagt caattttaca tgccataaa gacatttgta 3480
aagaattcag ctcttatggg ctgttgata aatgtgtatc taggcacttt attataactg 3540
gaatttgacc tcacagatgt tacaactga tcagtcgttt gacctaatgt gtggtagcta 3600
tctgtatgtt ttgcaatctt aatacagaca tgctttccaa aaagattaat acagaacct 3660
cctgccgttt tggataagtc tatccagctg tggaaagggc aacctgtggg ttctctgtac 3720
tgggtgttaa tgggggaaga atatgaacag ctttaaagag ctgtgtattg tggttactac 3780
tattaaaaaa taagatctgc acgagtctga ctggcctttg ggtggccttt gtggacggct 3840
cgtagctgga aagtgtgat ctgggttttc tggcattctt ttaagttaaa aagttaacat 3900
cgggacatgg gtttgatctt ttgttgatcc tgatgacagt gcagagattc tccacagctg 3960
gataaaaatg tcacaaagct acttactgta catgggcagt atcagatttc aaatccta at 4020
atctcagctg tgctttta at actcaaaaata ttaggggatg ggggtgtgaa gctttccctt 4080

100

ttttgctttt aacaatttat agaatttaac agatgtactg tttttcatgt ggcctcacat 4140
 ttaaagttat gagaacatac acatggttta caactttttac tatatacctt tccttggcca 4200
 ccaagtattt taaaagtgtg ccacctttta acctttactt tttttaagtt gaaggtgata 4260
 ctttttctat atatgatgaa actcatgtca actgaagtga gtgtaatctc agatatcaac 4320
 attattatat tttaaaatca cgctatggaa atatcacctg aattctgtca tttgtcagat 4380
 ttacagtacc tttttttctt taacttttag cattaaataa aaataaaatt gggagcactg 4440
 aa 4442

<210> 69
 <211> 2056
 <212> DNA
 <213> Homo sapien

<400> 69
 cccgcccgcg cggtcgtcgg agcgcggtc agcggcgcgg cggagactcg gcacggcggc 60
 ggccaggcgc aggcgggcgg gcggagcagc ggacggcgcc gaggcgcgcc cacgcgcggg 120
 gctcgcgggc gcgcagcgcg ccggagtggg cggggcccgc ggccgctcgc gcctctcgat 180
 gggcagctcg cacttgctca acaagggcct gccgcttgta agactttgac cccgacgcag 240
 gcttcagtga ttgccgcgga tcagaagcct gagcactttg gctcgagaca attaaggacg 300
 tgggatgagg ctccgagaca ggacgcggtt ctgcctgggg atcctgaaga taaaagcctt 360
 tgaaaagtcg aattcatggt cgtggaagct gagcccatat taagagatgt caggcgctccg 420
 acctccgac atgaacgggc ccctgcaccc gcggcccctg gtggcattgc tggatggccg 480
 ggactgcaca gtggagatgc ccatcctgaa ggacgtggcc actgtggcct tctgcgacgc 540
 gcagtccacg caggagatcc atgagaaggt cctgaacgag gctgtggggg ccctgatgta 600
 ccacaccatc actctcacca gggaggacct ggagaagtgc aaagccctcc gcatcatcgt 660
 ccggattggc agtggttttg acaacatcga catcaagtcg gccggggatt taggcattgc 720
 cgtctgcaac gtgcccgcgg cgtctgtgga ggagacggcc gactcgacgc tgtgccacat 780
 cctgaacctg taccggcggg ccacctggct gcaccaggcg ctgcgggagg gcacacgagt 840
 ccagagcgtc gagcagatcc gcgaggtggc gtccggcgct gccaggatcc gcggggagac 900
 cttgggcac atcggacttg gtcgcgtggg gcaggcagtg gcgctgcggg ccaaggcctt 960
 cggcttcaac gtgctcttct acgaccctta cttgtcggat ggctggagc gggcgtggg 1020
 gctgcagcgt gtcagcacc tgcaggacct gctcttccac agcgactgcg tgaccctgca 1080
 ctgcggcctc aacgagcaca accaccacct catcaacgac ttcaccgtca agcagatgag 1140
 acaaggggccc ttcttggtga acacagcccg ggggtggcctg gtggatgaga aggcgtggc 1200

101

ccaggccctg aaggagggcc ggatccgcgg cgcggccctg gatgtgcacg agtcggaacc 1260
 cttcagcttt agccagggcc ctctgaagga tgcacccaac ctcatctgca ccccccattgc 1320
 tgcattgtac agcgagcagg catccatcga gatgcgagag gaggcggcac gggagatccg 1380
 cagagccatc acaggccgga tcccagacag cctgaagaac ttctgccctg tgtccttcgc 1440
 gttcctcggt aagcagaaga agtcagtagt tattctccca tgaacgttct tgtctgtgta 1500
 cagttttttag aacattacaa aggatctgtt tgcttagctg tcaacaaaaa gaaaacctga 1560
 aggagcattt ggaagtcaat ttgaggtttt ttttttttgg tttttttttt ttgtatgttg 1620
 gaacgtgccc cagaatgagg cagttggcaa acttctcagg acaatgaatc cttcccgttt 1680
 ttctttttat gccacacagt gcattgtttt ttctacctgc ttgtcttatt tttagaataa 1740
 tttagaaaaa caaaacaaag gctgtttttc ctaatttttg catgaacccc cccttgttcc 1800
 aaatgaagac ggcatcacga agcagctcca aaaggaaaag cttggggcgg gcccagcgtg 1860
 cccgctgccc atcgacgtct gtcttgggga cgtggagggt ggcagcgtcc ccgctgcac 1920
 cagtgccgtc ctgctgatgt ggtaggctag caatatcttg gttaaaatca tgtttgtgac 1980
 tgtaaccatt tgtatgaatt attttaaaga aataaaaatc ctggaaagag ccagcgtgcc 2040
 cagcaaactc attctg 2056

<210> 70
 <211> 1862
 <212> DNA
 <213> Homo sapien

<400> 70
 cagtggggga gtttgccgcc cgctgcgggg tgaggatgga gccacgggct ctaggcctgc 60
 cttgcctcgg tcctgtgagg ggctgctcaa cctcctcagc ctttcttttg gcagggctcc 120
 tgaccaacca ggagagagtg agggcccagt gtctgtgagt ggtcccatg ggccctggcc 180
 atggagtcac ggcatcccgt ccagacctgc agccccata gcaccttggg accccaggca 240
 gccctgggtt ggatgttcag ccgcaagagg agacaccccc acaggggcag tatcagcctg 300
 cagctcctgg ggctacagac cttttggctg gaaggggcca ggcagcatgt ccccccata 360
 gggctcctcc cacgagggac ttggagatta aatcccttgg gctccacac ccgccgtgt 420
 caggagcccc tgggtgtgtc gatggtcccg gggctgtgct cttgagctca gccagcctgc 480
 cgagcagagc ggggccgtgg ggtctgtggt tcccaggcag ggccccccac aggggggttcc 540
 agtgcacagc cacctcccct ccgtaccag ccacagcact caggctgcac tgaccatgcc 600
 tgtgtgttcc ccagctttag ccagggccct ctgaaggatg caccacaact catctgcacc 660
 ccccatgctg catggtacag cgagcaggca tccatcgaga tgcgagagga ggcggcacgg 720

102

gagatccgca gagccatcac aggccggatc ccagacagcc tgaagaactg tgtcaacaag 780
 gaccatctga cagccgccac ccaactgggcc agcatggacc ccgccgtcgt gcaccctgag 840
 ctcaatgggg ctgcctatag gtaccctccg ggcgtggtgg gcgtggcccc cactggcatc 900
 ccagctgctg tggaaggat cgtccccagc gccatgtccc tgtcccacgg cctgccccct 960
 gtggcccacc cgccccacgc cccttctcct ggccaaaccg tcaagcccga ggcggataga 1020
 gaccacgcca gtgaccagtt gtagcccggt aggagctctc cagcctcggc gcctgggcag 1080
 agggcccggg aaccctcgga ccagagtgtg tggaggaggc atctgtgtgg tggccctggc 1140
 actgcagaga ctgggtccggg ctgtcaggag gcgggagggg gcagcgctgg gcctcgtgtc 1200
 gcttgctcgtc gtccgtcctg tgggcgtctt gccctgtgtc cttcgcgttc ctcgtaaagc 1260
 agaagaagtc agtagttatt ctcccatgaa cgttcttgtc tgtgtacagt ttttagaaca 1320
 ttacaaagga tctgtttgct tagctgtcaa caaaaagaaa acctgaagga gcatttgga 1380
 gtcaatttga ggtttttttt tttgggtttt ttttttttgt atgttggaac gtgccccaga 1440
 atgaggcagt tggcaaaact ctcaggacaa tgaatccttc ccgtttttct ttttatgcca 1500
 cacagtgcac tgttttttct acctgcttgt cttattttta gaataattta gaaaaacaaa 1560
 acaaaggctg tttttcctaa ttttggcatg aacccccctt tgttccaaat gaagacggca 1620
 tcacgaagca gtcctaaaag gaaaagcttg ggcgggtgcc agcgtgcccg ctgcccacg 1680
 acgtctgtcc tggggacgtg gaggggtggc gcgtccccgc ctgcaccagt gccgtcctgc 1740
 tgatgtggta ggctagcaat attttgggta aaatcatgtt tgtgactgta accatttgta 1800
 tgaattattt taaagaaata aaaatcctgg aaagagccag cgtgcccagc aaatctattc 1860
 tg 1862

<210> 71
 <211> 1168
 <212> DNA
 <213> Homo sapien

<400> 71
 ctgcatggta cagcgagcag gcatccatcg agatgcgaga kgaggcacc ttcagcttta 60
 gccagggccc tctgaaggat gcacccaacc tcatctgcac ccccatgct gcatggtaca 120
 tggacccgc cgctgtgcac cctgagctca atggggctgc ctataggtag cctccgggag 180
 tgggtggcgt ggccccact ggcacccag ctgctgtgga aggtatcgt cccagcgcca 240
 tgctcctgtc ccacggcctg cccctgtgg cccacccgcc ccacgcccct tctcctggcc 300
 aaaccgtcaa gcccgaggcg gatagagacc acgccagtga ccagttgtag cccgggagga 360
 gctctccagc ctcggcgcct gggcagaggg cccggaaacc ctcggaccag agtgtgtgga 420

103

```

ggaggcatct gtgtggtggc cctggcactg cagagactgg tccgggctgt caggaggcgg 480
gagggggcag cgctgggcct cgtgtcgctt gtcgtcgccc gtccctgtggg cgctctgccc 540
tgtgtccttc gcgttcctcg ttaagcagaa gaagtcagta gttattctcc catgaacgtt 600
cttgtctgtg tacagttttt agaacattac aaaggatctg tttgcttagc tgtcaacaaa 660
aagaaaacct gaaggagcat ttggaagtca atttgagggt tttttttttg gttttttttt 720
ttttgtatgt tggaacgtgc ccagaaatga ggcagttggc aaacttctca ggacaatgaa 780
tccttcccggt ttttcttttt atgccacaca gtgcattgtt ttttctacct gcttgtctta 840
tttttagaat aatttagaaa aacaaaacaa aggctgtttt tcctaatttt ggcatgaacc 900
cccccttggt ccaaatgaag acggcatcac gaagcagctc caaaaggaaa agcttgggcg 960
gtgcccagcg tgcccgtgc ccacgcagct ctgtcctggg gacgtggagg gtggcagcgt 1020
ccccgcctgc accagtgcg tctgtctgat gtggtaggct agcaatattt tgggttaaat 1080
catgtttgtg actgtaacca tttgtatgaa ttatttttaa gaaataaaaa tcctggaaag 1140
agccagcgtg ccagcaaat ctattctg 1168

```

```

<210> 72
<211> 1352
<212> DNA
<213> Homo sapien

```

```

<400> 72
gggcatcttg tggggtggga cggcttgtgg ggtggggcat ctccggagggt ggggcatctc 60
tggggcccg ccacttggga ggcggggcat cctggggcg gggcatctca gagggcgctt 120
ccggaggctg gagtatcttg ggaggtggga gcagggtggca gagaggcttc ccacagctgg 180
ctggaggcgt gatcctgggt gtggccctgt ggctccgcca tgaccgcag accaccaacc 240
tcctgtatct ggagctggga gacaagcccg cgcccaacac cttctatgta ggcatctaca 300
tcctcatcgc tgtgggcgct gtcattgatgt tcgttggctt cctgggctgc tacggggcca 360
tcagggaatc ccagtgcctg ctggggacgt tcttcacctg cctggtcac ctgtttgcct 420
gtgaggtggc cgcgggcac tggggctttg tcaacaagga ccagatcgcc aaggatgtga 480
agcagttcta tgaccaggcc ctacagcagg ccgtggtgga tgatgacgcc aacaacgcca 540
aggctgtggt gaagaccttc cagcagacgc ttgactgctg tggctccagc aactgactg 600
ctttgaccac ctcatgtctc aagaacaatt tgtgtccctc gggcagcaac atcatcagca 660
acctcttcaa ggaggactgc caccagaaga tcgatgacct cttctccggg aagctgtacc 720
tcacggcat tgctgccac gtggtcgctg tgatcatgat cttcgagatg atcctgagca 780
tggtgctgtg ctgtggcac cggaacagct ccgtgtactg agggcccgca gctctggcca 840

```

104

cagggacctc tgcagtgcc cctaagtgac cgggacactt ccgagggggc catcaccgcc 900
 tgtgtatata acgtttccgg tattactctg ctacacgtag cctttttact tttggggttt 960
 tgtttttgtt ctgaactttc ctgttacctt ttcagggctg acgtcacatg taggtggcgt 1020
 gtatgagtgg agacgggcct ggtctttggg gactggaggg caggggtcct tctgccctgg 1080
 ggtcccaggg tgctctgcct gctcagccag gcctctcctg ggagccactc gcccagagac 1140
 tcagcttggc caacttgggg ggctgtgtcc acccagcccg cccgtcctgt gggtgcaca 1200
 gctcaccttg ttccctcctg ccccggttcg agagccgagt ctgtgggca cctctgcctt 1260
 catgcacctg tcctttctaa cacgtgcct tcaactgtaa tcacaacatc ctgactccgt 1320
 catttaataa agaaggaaca tcaggcatgc ta 1352

<210> 73
 <211> 1445
 <212> DNA
 <213> Homo sapien

<400> 73
 agatgagtgg ggcagtcaca tcccaccttc cccaagccgg gctgttctgc acagcctgct 60
 tgggacgctg gtgggagtca ctgtggcctt cggcactgcc ctggcagtgg gggcagctag 120
 gccatttggg aggggctcgc ctccccagg cccggccctg ggacctcagc cgttgcttag 180
 tgggtggcctg cttcagccca ggcattgtgg agaggcacca gacacaggat gtcctctgc 240
 cagcccctga agccccgtcc cctgacgagc tggctggagg cgtgatcctg ggtgtggccc 300
 tgtggctccg ccatgaccgg cagaccacca acctcctgta tctggagctg ggagacaagc 360
 ccgcgccc aa cactttctat gtaggcatct acatcctcat cgctgtgggc gctgtcatga 420
 tgttcgttgg cttcctgggc tgctacgggg ccatccagga atcccagtgc ctgctgggga 480
 cgttcttcac ctgcctggtc atcctgtttg cctgtgaggt ggccgccggc atctggggct 540
 ttgtcaacaa ggaccagatc gccaaaggatg tgaagcagtt ctatgaccag gccctacagc 600
 aggccgtggt ggatgatgac gccacaacg ccaaggctgt ggtgaagacc ttccacgaga 660
 cgcttgactg ctgtggctcc agcacactga ctgctttgac cacctcagtg ctcaagaaca 720
 atttgtgtcc ctccggcagc aacatcatca gcaacctctt caaggaggac tgccaccaga 780
 agatcgatga cctcttctcc gggaagctgt acctcatcgg cattgctgcc atcgtggtcg 840
 ctgtgatcat gatcttcgag atgatcctga gcatggtgct gtgctgtggc atccggaaca 900
 gctccgtgta ctgaggcccc gcagctctgg ccacagggac ctctgcagtg ccccctaagt 960
 gacccggaca cttccgaggg ggccatcacc gcctgtgtat ataacgtttc cggtattact 1020
 ctgctacacg tagccttttt acttttgggg ttttgttttt gttctgaact ttctgttac 1080

105

cttttcaggg ctgacgtcac atgtaggtgg cgtgtatgag tggagacggg cctgggtctt 1140
 ggggactgga gggcaggggt ccttctgccc tgggggtccca ggggtgctctg cctgctcagc 1200
 caggcctctc ctgggagcca ctgcccaga gactcagctt ggccaacttg gggggctgtg 1260
 tccaccacgc ccgcccgtcc tgtgggctgc acagctcacc ttgttccttc ctgccccggt 1320
 tcgagagccg agtctgtggg cactctctgc cttcatgcac ctgtcctttc taacacgtcg 1380
 ccttcaactg taatcacaac atcctgactc cgtcatttaa taaagaagga acatcaggca 1440
 tgcta 1445

<210> 74
 <211> 2290
 <212> DNA
 <213> Homo sapien

<400> 74
 gtcaggacca aaaaaggcaa cgcgcctctc ctgacctgta ccccgagtg aacccaacct 60
 tgcaaccacag gagtgtcagg gcctgagggg agggagacct ggctcctggg tgccgtgccc 120
 gtaaggaggt ggccacctgc agggcattcc tggcagaggc ttcatctggc caggtaggag 180
 gctgggtggc cgagcccca atctgggtgt gttctctgcc tggcggtggg tcctgcccc 240
 ggcaccttct cctctgggct ggctgggcag ggacaatggg cctggctgcg aggagggggc 300
 ctgggctgcc ttctgcattg cctcggtgac gggagatggc ccctgcctgc tgagggatag 360
 gggagtgggc aggcagttag agacactgac agctgtcccc cgggtacagg gcctgtctg 420
 ggtggccagg cccatgtctc gggcccacag tgcgcccccc acccttggac ggcgccttct 480
 ccctccccag gtgcatgctg ccagccagg gagcgtgggg gagttcggga gggctggcct 540
 acacgccctg gtccagctgt ccaggtggg gtgctgggct tcagccctca gccagggcc 600
 taggaatcca acttgatcct cccacacag cagccaggtt caaatgcagg tcccgtaacg 660
 gaagtgtgc tgtgcagccc agattggggg gcaggagcca gcagggcccc ccacacctct 720
 tctcgacca cactggggag gcagcattgg ttccagttcc ggttcctggg ctgccctctc 780
 aaccccgcc tacagtggg cccacctgt gccttctgat gccactcca cccacgcca 840
 agtcccagag gctttgggag cgggtgaagg cgggtgggtg cgggtggcag gtgcaggcgg 900
 tgggtggtgg gtgtggcagg tggcgggccc caccgcaggt gtcacccctg cgaagcacct 960
 gtcgccagca ctgagagcgc tcatgaggtg ccagtcctcc atgtggcctc cttagtctcc 1020
 gtccgtgtgc atggaagagg taactgaggc acagaaaact caccaggcca ggctgggatg 1080
 tgaggctcct tgctgctcat ccctggcagt cagcaaccct acatcttccc agctgggcgg 1140
 cccgtggtgg gttcggcacc caggacctc cggggctctg ggctgtggcg agtgtgtagg 1200

106

caccacactg gtgtctctct ccccgcaagg catctacatc ctcacgctg tgggcgctgt 1260
 catgatgttc gttggcttcc tgggctgcta cggggccatc caggaatccc agtgcctgct 1320
 ggggacgttc ttcacctgcc tgggtcatcct gtttgctgt gaggtggccg ccggcatctg 1380
 gggctttgtc aacaaggacc agatcgccaa ggatgtgaag cagttctatg accaggccct 1440
 acagcaggcc gtggtggatg atgacgccaa caacgccaag gctgtggtga agacctcca 1500
 cgagacgctt gactgctgtg gctccagcac actgactgct ttgaccacct cagtgtcaa 1560
 gaacaatttg tgtccctcgg gcagcaacat catcagcaac ctcttcaagg aggactgcca 1620
 ccagaagatc gatgacctct tctccgggaa gctgtacctc atcggcattg ctgccatcgt 1680
 ggtcgctgtg atcatgatct tcgagatgat cctgagcatg gtgctgtgct gtggcatccg 1740
 gaacagctcc gtgtactgag gccccgcagc tctggccaca gggacctctg cagtgcctcc 1800
 taagtgacct ggacacttcc gagggggcca tcaccgctg tgtatataac gtttccggtg 1860
 ttactctgct acacgtagcc tttttacttt tggggttttg tttttgttct gaactttcct 1920
 gttacctttt cagggtgac gtcacatgta ggtggcgtgt atgagtggag acgggcctgg 1980
 gtcttgggga ctggagggca ggggtccttc tgcctgggg tcccagggtg ctctgcctgc 2040
 tcagccaggc ctctcctggg agccactcgc ccagagactc agcttggcca acttgggggg 2100
 ctgtgtccac ccagcccgcc cgtcctgtgg gctgcacagc tcacctgtt cctcctgccc 2160
 ccggttcgag agccgagtct gtgggcactc tctgccttca tgcacctgtc ctttctaaca 2220
 cgtcgccttc aactgtaatc acaacatcct gactccgtca ttaataaag aaggaacatc 2280
 aggcattgcta 2290

<210> 75
 <211> 1033
 <212> DNA
 <213> Homo sapien

<400> 75
 ccattgtgct ggaaagagcg cgcaacggcg gcgacggcgg cgacccacc gcgcactcctg 60
 ccaggcctcc gcgccagcc gccacgcgc ccccgcgccc cgcgcccga ccctttcttc 120
 gcgccccgc ccctcgccc gccaggcccc cttgcgggcc acccgccagg ccccgcgccg 180
 gccccccgc cgccaggac cggccccgc cccgcaggcc gcccgccgc cgcgcgcca 240
 tgggagtgga gggtgcacc aagtgcata agtacctgct ctctgtcttc aatttcgtct 300
 tctggctggc tggaggcgtg atcctgggtg tggccctgtg gctccgcat gaccgcaga 360
 ccaccaacct cctgtatctg gagctgggag acaagccgc gcccaacacc ttctatgtag 420
 gcatctacat cctcatcgct gtgggcgctg tcatgatgtt cgttggcttc ctgggctgct 480

107

```

acggggccat ccaggaatcc cagtgcctgc tggggacgtt cttcacctgc ctggtcaccc 540
tgtttgccctg tgaggtggcc gccggcatct ggggctttgt caacaaggac cagatcgcca 600
aggatgtgaa gcagttctat gaccaggccc tacagcaggc cgtggtggat gatgacgcca 660
acaacgcaa ggctgtggtg aagaccttcc acgagacgct tgactgctgt ggctccagca 720
cactgactgc tttgaccacc tcagtgtca agaacaattt gtgtccctcg ggcagcaaca 780
tcatacagcaa cctcttcaag gaggactgcc accagaagat cgatgacctc ttctccggga 840
agctgtacct catcggcatt gctgccatcg tggtcgctgt gatcatgata ttcgagatga 900
tcttgagcat ggtgctgtgc tgtggcatcc ggaacagctc cgtgtactga ggccccgcag 960
ctctggccac agggacctct gcagtgcccc ctaagtgacc cggacacaat gttcaggaca 1020
cgaatgttca ggc 1033

```

```

<210> 76
<211> 1190
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1081)..(1081)
<223> n=a, c, g or t

```

```

<220>
<221> misc_feature
<222> (1122)..(1122)
<223> n=a, c, g or t

```

```

<400> 76
ccattgtgct ggaaagagcg cgcaacggcg gcgacggcgg cgaccccacc gcgcatcctg 60
ccaggcctcc gcgcccagcc gcccaacgcgc ccccgcgccc cgcgccccga ccctttcttc 120
gcgccccgc ccctcggccc gccaggcccc cttgccggcc acccgccagg ccccgcgccg 180
gcccgccgc cgcccaggac cggcccgcg cccgcaggcc gcccgccgc cgcgccgcca 240
tgggagtgga gggctgcacc aagtgcata agtacctgct cttcgtcttc aatttcgtct 300
tctggctggc tggaggcgtg atcctgggtg tggccctgtg gctccgcat gaccgcaga 360
ccaccaacct cctgtatctg gagctgggag acaagccgc gcccaacacc ttctatgtag 420
gcatacatat cctcatcgct gtgggcgctg tcatgatgtt cgttggcttc ctgggctgct 480
acggggccat ccaggaatcc cagtgcctgc tggggacgtt cttcacctgc ctggtcaccc 540
tgtttgccctg tgaggtggcc gccggcatct ggggctttgt caacaaggac cagatcgcca 600
aggatgtgaa gcagttctat gaccaggccc tacagcaggc cgtggtggat gatgacgcca 660

```

108

acaacgccaa ggctgtggtg aagaccttcc acgagacgct tgactgctgt ggctccagca 720
 cactgactgc ttgaccacc tcagtgtca agaacaattt gtgtccctcg ggcagcaaca 780
 tcacagcaa cctcttcaag gaggactgcc accagaagat cgatgacctc ttctccggga 840
 agctgtacct catcggcatt gctgccatcg tggtcgctgt gatcatgac ttcgagatga 900
 tcctgagcat ggtgctcaat gataatcttt gtattatagg aaaagtaagg attagtggaa 960
 gacagggttt ctatccaaac caacaacata aacggcaata taattgttaa aattagcctt 1020
 tatatgccat atgttatgcc tggtagatag tagactccca ataaatcttt ggtggcttga 1080
 nttgaagagg taaagtccag atgaaatctt tggcctcctt tncctggccg ggacagtaag 1140
 attctcatgc taagcaataa aatggtctta tctcttaata atccccatcc 1190

<210> 77
 <211> 871
 <212> DNA
 <213> Homo sapien

<400> 77
 ccattgtgct ggaaagagcg cgcaacggcg gcgacggcgg cgacccacc gcgcacctg 60
 ccaggcctcc gcgcccagcc gccacgcgc ccccgcgccc cgcgcccga cctttcttc 120
 gcgccccgc cctcggccc gccaggcccc ctgcccggcc acccgccagg ccccgcgccg 180
 gcccgcccgc cgcccaggac cggcccgcgc ccgcaggcc gcccgccgc cgcgccgcca 240
 tgggagtgga gggctgcacc aagtgcac caagctgct ctctgtcttc aatttcgtct 300
 tctggctggc tggaggcgtg atcctgggtg tggccctgtg gctccgcat gaccgcaga 360
 ccaccaacct cctgtatctg gagctgggag acaagcccgc gcccaacacc ttctatgtag 420
 gcatctacat cctcatcgct gtgggcgtg tcatgatgtt cgttggcttc ctgggctgct 480
 acggggccat ccaggaatcc cagtgcctgc tggggacgtt cttcacctgc ctggtcatcc 540
 tgtttgcctg tgaggctggc gccggcatct ggggctttgt caacaaggac cagatcgcca 600
 aggatgtgaa gcagttctat gaccaggccc tacagcaggc cgtggtggat gatgacgcca 660
 acaacgccaa ggctgtggtg aagaccttcc acgagacgct tgactgctgt ggctccagca 720
 cactgactgc ttgaccacc tcagtgtca agaacaattt gtgtccctcg ggcagcaaca 780
 tcacagcaa cctcttcaag gaggactgcc accagaagat cgatgacctc ttctccggga 840
 agctgtacct ggccgcgacc acgctaaggg c 871

<210> 78
 <211> 1283
 <212> DNA
 <213> Homo sapien

109

<400> 78
 ggaatcacat cgagccactg aaaattcagt ggctcgatgt gcttcagcga gaaccagac 60
 ctttcccaaa gtcaggatt cttcgaaaag ttgagaaaat tgatgacttc aaagctgaag 120
 actttcagat tgaagggtac aatccgcac caactattaa aatggaaatg gctgtttagg 180
 gtgctttcaa aggagctcga aggatattgt cagtcttttag gggttgggct ggatgccgag 240
 gtaaaagtgc tttttgctct aaaagaaaaa ggaactaggt caaaaatctg tccgtgacct 300
 atcagttatt aatttttaag gatgttgcca ctggcaaatg taactgtgcc agttctttcc 360
 ataataaaag gctttgagtt aactcactga gggatatctga caatgctgag gttatgaaca 420
 aagtgaggag aatgaaatgt atgtgctctt agcaaaaaca tgtatgtgca tttcaatccc 480
 acgtacttat aaagaagggt ggtgaatttc acaagctatt tttggaatat ttttagaata 540
 ttttaagaat ttcacaagct attccctcaa atctgaggga gctgagtaac accatcgatc 600
 atgatgtaga gtgtgggtat gaactttaaa gttatagttg ttttatatgt tgctataata 660
 aagaagtgtt ctgcattcgt ccacgctttg ttcattctgt actgccactt atctgctcag 720
 ttccttccta aaatagatta aagaactctc cttaagtaaa catgtgctgt attctggttt 780
 ggatgctact taaaagagta ttttttagaa ataatagta atatattttg ccctattttt 840
 ctcattttaa ctgcattcta tcctcaaaat ataatgacca tttaggatag agtttttttt 900
 tttttttttt taaactttta taaccttaaa gggttatttt aaaataatct atggactacc 960
 attttgcct cattagcttc agcatgggtg gacttctcta ataatatgct tagattaagc 1020
 aaggaaaaga tgcaaaacca cttcgggggt aatcagtga atatttttcc cttcgttgca 1080
 taccagatac ccccggtgtt gcacgactat ttttattctg ctaatttatg acaagtgtta 1140
 aacagaacaa ggaattattc caacaagtta tgcaacatgt tgcttatttt caaattacag 1200
 tttaatgtct aggtgccagc ccttgatata gctatttttg taagaacatc ctctgggact 1260
 ttgggttagt taaatcgctg act 1283

<210> 79
 <211> 1169
 <212> DNA
 <213> Homo sapien

<400> 79
 ttcatttctt aattggcatt tagtttagca tagtaatttt ttttttttaa tttgctgcca 60
 cttcagcagt tttagtgttt agatgtggga gaggttaagt aagatctgta ttgccctatt 120
 atctctaaaa gtggcttgct cctgcaggat agtgaatgct tgtctgtacc attggagtta 180
 cacatttcat tgagttgata ccagcagatt tttttcaggg aaaaataatc caatctcata 240
 actacagcac ctcaccacat tctccctctc ctaacagtgt agagacttag atttaattgc 300

```

acctttcctt tcaatagtta tttctggttg gatgttgaag aaagctagtg ggcaaaggat 360
tccgcttgaa cagacactgg tgatgggctc ccaggcttgg cggtggttg ctttctgcac 420
tgcagtaaaa ccacgagagg gcatcttttg ctctgggatt aataattcat ctgtcttctc 480
tgacctgtga ctttttctct cttttctctt accctttccc ggtagtgtga agtgaggggg 540
tctctctccc tcttctctct tctctgtga ttcaccttc tttttaccct gccctgcggc 600
ggctccgccc cttaccttca tggacgactc agaggtggag tcgaccgcca gcatcttggc 660
ctctgtgaag gaacaagagg ccagtttga gaagctgacc cgggcgctgg aggaggaacg 720
gcgccacgtc tcggcgcgagc tggacgcgt ccgggtctca ccacaagatg ccaaccact 780
catggccaac ggcacatccc cttttcgaaa gaaatgcaaa aaaaagagta tttttcaag 840
ccgtgtcgaa ctatttaagg aaagcaaat catttcctaa atgcatatca tttgtgagaa 900
tttctcatta atatctgaa tcattcattt cagctaaggc ttcattgtga ctcgatatgt 960
catctaggaa agtactattt catggtccaa acctgttgcc atagttggta aggctttcct 1020
ttaagtgtga aatatttaca tgaaattttc tcttttaaag ttctttatag ggttaggggtg 1080
tgggaaaatg ctatattaat aaatctgtag tgtttagtgt ttatatgttc agaaccagag 1140
tagactggat tgaaagatgg actgggtct 1169

```

```

<210> 80
<211> 406
<212> DNA
<213> Homo sapien

```

```

<400> 80
gatgatcata tattagcatg gttcataatc atgtcgacgc gcagtgtatg atattgcaga 60
atcggcttac gtgtcgcgcg agtcctccgg ggtatggcca tgtacaccag caaggaccgc 120
tatttctatt ttgggaagct tgatggccag atctcctctg cctaccccag ccaagagggg 180
cagggtgctgg tgggcatcta tggccagtat caactccttg gcatcaagag cattggcttt 240
gaatggaatt atccactaga ggagccgacc actgagccac cagttaatct cacatactca 300
gcaaactcac ccgtgggtcg ctaggggtgg gtatggggcc atccgagctg aggccatctg 360
ggtgggtgggtg gctgatggta cctgcccggc ggcgctcgaa acgccc 406

```

```

<210> 81
<211> 1902
<212> DNA
<213> Homo sapien

```

```

<400> 81
tataggcgcc tgggtttcta atctgtcgag cggcgagtg tgatggatga gcggccgccc 60

```

111

gggcaggtgg cgagagccct gggatgcacc ggccagaggc catgctgctg ctgctcacgc	120
ttgccctcct ggggggcccc acctgggcag ggaagatgta tggccctgga ggaggcaagt	180
atttcagcac cactgaagac tacgaccatg aaatcacagg gctgcgggtg tctgtaggtc	240
ttctcctggt gaaaagtgtc caggtgaaac ttggagactc ctgggacgtg aaactgggag	300
ccttaggtgg gaatacccgag gaagtcaccc tgcagccagg cgaatacatc acaaaagtct	360
ttgtcgctt ccaagctttc ctccggggtg tggatcatgta caccagcaag gaccgctatt	420
tctatttttg gaagcttgat ggccagatct cctctgcta cccagccaa gaggggcagg	480
tgctggtggg catctatggc cagtatcaac tccttgcat caagagcatt ggctttgaat	540
ggaattatcc actagaggag ccgaccactg agccaccagt taatctcaca tactcagcaa	600
actcaccctg gggtcgctag ggtggggtat ggggccatcc gagctgaggc catctgggtg	660
gtgggtggctg atggtactgg agtaactgag tggggacgct gaatctgaat ccaccaataa	720
ataaagcttc tgcagaatca gtgcatccag gattggctct tggatctggg gtacaaccaa	780
agccttcctt gtccttgga gacaaagtcc ccagtgtgc agcccagtga actgagatga	840
ggggtagggc aaaggtgact ctgccgagga cagaaagaga gcagcaccac cccctcagag	900
gtgctgtgga tctctgtgcc agtcccacaa tctttgaaga gtcaggcttc aaggccgcca	960
ctccccactg tccctcacc caggcccacc cagcgggct ctgctggcca ctcagtcctg	1020
ggattctctg agtgctcaag gggcctcagg gaagccactc actcatccat tctcagca	1080
aacatttgcc aaggccgtgt cctcaccagg ctgcctggc actgggggta taaagaagag	1140
gccaggtttc agctgagtgc gtatacactg tgggtggggc tgggtgcagag gcagatagt	1200
gcagttcact ttgccatctg catggatggg gacacatgca ggtcatttg cgtaggcacc	1260
tgcccagagg tgaggtgaga gggatgtgct tcagggaagt ctctctagag gaggcaatgt	1320
ctgagctaag tcttaaagga tgaaggacaa ttggccaagt ggagacaaag ggaggagggg	1380
attacaggtt gagggctcaa catgagcaaa gacaggagat gcagatggct ggggtaattt	1440
gtcacagtag caacaggaaa tcagtacaag ccctggacct ggcccattct tcatgtcccc	1500
ttcccaagcc tccaagccca catgggcact tgccaagatc agagctccag gggcctccag	1560
ggacgggggtt ttcagtcctt tgggaccatg atccacagag agaaattgat cctacttgag	1620
acacaggaag caacacatgt aatgcaacag ccaagcacc aacgatcatc taatgcaaca	1680
gccaaacacc cagcgatcat ctaatgcaac agccaaacac ccagcgatca tctaatagcaa	1740
cagccaaaca cccagcgatc atctaatagca acagccaaac acccagcgat catctaatagc	1800
aacagccaaa caccagtgta tcatctaata caacagccaa acaccagtg atcatctaata	1860
gcaacagcca aacagacaag tgatcatcta atgcaacagc ca	1902

112

<210> 82
<211> 1911
<212> DNA
<213> Homo sapien

<400> 82
catgccgagc ggcgcagtgat gatggatcgg ccgccgggca ggtggaggggt gcccggcaca 60
accagacgcc cagtcacggc gagagccctg ggatgcaccg gccagaggcc atgctgctgc 120
tgctcacgct tgccctcctg gggggcccca cctgggcagg gaagatgtat ggccctggag 180
gaggcaagta ttccagcacc actgaagact acgaccatga aatcacaggg ctgcggggtgt 240
ctgtagggtct tctcctgggtg aaaagtgtcc aggtgaaact tggagactcc tgggacgtga 300
aactgggagc cttagggtggg aatacccagg aagtcaccct gcagccaggc gaatacatca 360
caaaagtctt tgtcgcttcc caagctttcc tccgggggat ggtcatgtac accagcaagg 420
accgctatct ctatcttggg aagcttgatg gccagatctc ctctgcctac cccagccaag 480
aggggcagggt gctgggtggc atctatggcc agtatcaact ccttggcatc aagagcattg 540
gctttgaatg gaattatcca ctagaggagc cgaccactga gccaccagtt aatctcacat 600
actcagcaaa ctcacccgtg ggtcgctagg gtgggggatg gggccatccg agctgaggcc 660
atctgggtgg tgggtggctga tgggtactgga gtaactgagt cgggacgctg aatctgaatc 720
caccaataaa taaagcttct gcagaatcag tgcattccagg attggtcctt ggatctgggg 780
tacaacaaa gccttcctg ctcttggag acaaagtccc cagtgtgca gcccagtga 840
ctgagatgag gggtagggca aagggtgactc tgccgaggac agaaagagag cagcaccacc 900
ccctcagagg tgctgtggat ctctgtgcca gtcccacaat ctttgaagag tcaggcttca 960
aggccgccac tccccactgt ccctcaccac agggccacc agcgggcctc tgctggccac 1020
tcagtcctcg gattctctga gtgtcaagg ggcctcagg aagccactca ctcatccatt 1080
cactcagcaa acatttgcca agggcgtgtc ctccaccagg tcgcctggca ctgggggtat 1140
aaagaagagg ccagggttca gctgagtgcg tatacactgt ggtgggggct ggtgcagagg 1200
cagatagtgg cagttcactt tgccatctgc atggatgggg acacatgcag gtcattgtgc 1260
gtaggcacct gccagagggt gaggtgagag ggatgtgctt cagggaagtc ttcctagagg 1320
aggcaatgtc tgagctaagt cttaaaggat gaaggacaat tggccaagt gagacaaagg 1380
gaggagggga ttacagggtg agggctcaac atgagcaaag acaggagatg cagatggctg 1440
gggtaatttg tcacagtagc aacaggaaat cagtacaagc cctggacctg gccattctt 1500
catgtcccct tcccaagcct ccaagccac atgggcactt gccaatgca gagctccagg 1560
ggcctccagg gacgggggtt tcagtccttt gggaccatga tccacagaga gaaattgatc 1620

113

ctacttgaga cacaggaagc aacacatgta atgcaacagc caagcaccca acgatcatct 1680
aatgcaacag ccaaacaccc agcgatcatc taatgcaaca gccaaacacc cagcgatcat 1740
ctaagtgaac agccaaacac ccagcgatca tctaagtcaa cagccaaaca cccagcgatc 1800
atctaagtga acagccaaac acccagtgat catctaagtc aacagccaaa caccagtgta 1860
tcatctaagt caacagccaa acagacaagt gatcatctaa tgcaacagcc a 1911

<210> 83
<211> 1852
<212> DNA
<213> Homo sapien

<400> 83
atgcatgtcg agcggcgagc tgtgatggat cgtggtcgag gcgaggtggc catgctgctg 60
ctgctcagcg ttgccctcct ggggggcccc acctgggcag ggaagatgta tggccctgga 120
ggaggcaagt atttcagcac cactgaagac tacgaccatg aaatcacagg gctgcgggtg 180
tctgtaggtc ttctcctggt gaaaagtgtc caggtgaaac ttggagactc ctgggacgtg 240
aaactgggag ccttaggtgg gaataccagc gaagtcaccc tgcagccagg cgaatacatc 300
acaaaagtct ttgtcgccct ccaagcttcc ctccggggta tggatcatgta caccagcaag 360
gaccgctatt tctattttgg gaagcttgat ggccagatct cctctgccta cccagccaa 420
gaggggcagg tgctgggtgg catctatggc cagtatcaac tccttggcat caagagcatt 480
ggctttgaat ggaattatcc actagaggag ccgaccactg agccaccagt taatctcaca 540
tactcagcaa actcaccctg gggtcgctag ggtgggggat ggggccatcc gagctgaggc 600
catctgggtg gtggtggctg atgggtactg agtaactgag tcgggacgct gaatctgaat 660
ccaccaataa ataaagcttc tgcagaatca gtgcatccag gattgggtcct tggatctggg 720
gtacaaccaa agccttcctt gctccttggg gacaaaagtc ccagtgtgtc agccagtgta 780
actgagatga ggggtagggc aaagggtgact ctgccgagga cagaaagaga gcagcaccac 840
cccctcagag gtgtgtgga tctctgtgcc agtcccacaa tctttgaaga gtcaggcttc 900
aaggccgcca ctccccactg tccctcacc caggcccacc cagcgggcct ctgctggcca 960
ctcagtcccg ggattctctg agtgctcaag gggcctcagg gaagccactc actcatccat 1020
tcactcagca aacatttgcc aaggccgtgt cctcaccagg ctgcctggc actgggggta 1080
taaagaagag gccaggtttc agctgagtgc gtatacactg tgggtggggc tgggtgcagag 1140
gcagatagtg gcagttcact ttgccatctg catggatggg gacacatgca ggtcattgtg 1200
cgtaggcacc tgcccagagg tgagggtgaga gggatgtgct tcaggggaagt ctctcctagag 1260
gaggcaatgt ctgagctaag tcttaaagga tgaaggacaa ttggccaagt ggagacaaaag 1320

114

ggaggagggg attacaggtt gagggctcaa catgagcaaa gacaggagat gcagatggct 1380
 ggggtaattt gtcacagtag caacaggaaa tcagtacaag ccctggacct ggcccattct 1440
 tcatgtcccc ttccaagcc tccaagccca catgggcact tgccaagatc agagctccag 1500
 gggcctccag ggacgggggtt ttcagtcctt tgggaccatg atccacagag agaaattgat 1560
 cctacttgag acacaggaag caacacatgt aatgcaacag ccaagcacc aacgatcatc 1620
 taatgcaaca gccaaacacc cagcgatcat ctaatgcaac agccaaacac ccagcgatca 1680
 tctaatagcaa cagccaaaca cccagcgatc atctaatagca acagccaaac acccagcgat 1740
 catctaatagc aacagccaaa caccagtgat tcatctaatag caacagccaa acaccagtg 1800
 atcatctaata gcaacagcca aacagacaag tgatcatcta atgcaacagc ca 1852

<210> 84
 <211> 1798
 <212> DNA
 <213> Homo sapien

<400> 84
 catgctcgag cggcgcaaktg tgatggatcg cccgggcagg tccagtcaca gatgtatggc 60
 cctggaggag gcaagtattt cagcaccact gaagactacg accatgaaat cacagggctg 120
 cgggtgtctg taggtcttct cctggtgaaa agtgtccagg tgaaacttgg agactcctgg 180
 gacgtgaaac tgggagcctt aggtgggaat acccaggaag tcaccctgca gccaggcgaa 240
 tacatcaciaa aagtctttgt cgccttccaa gctttcctcc ggggtatggg catgtacacc 300
 agcaaggacc gctatttcta ttttgggaag cttgatggcc agatctcctc tgcctacccc 360
 agccaagagg ggcagggtgct ggtgggcac tctggccagt atcaactcct tggcatcaag 420
 agcattggct ttgaatggaa ttatccacta gaggagccga ccaactgagcc accagttaat 480
 ctcacatact cagcaaactc acccgtgggt cgctaggggt ggggtatggg ccatccgagc 540
 tgaggccatc tgggtgggtg tggctgatgg tactggagta actgagtcgg gacgctgaat 600
 ctgaatccac caataaataa agcttctgca gaatcagtgc atccaggatt ggtccttgga 660
 tctgggggtac aaccaaagcc ttccctgctc cttggagaca aagtccccag tgctgcagcc 720
 cagtgaactg agatgagggg tagggcaaag gtgactctgc cgaggacaga aagagagcag 780
 caccaccccc tcagagggtgc tgtggatctc tgtgccagtc ccacaatctt tgaagagtca 840
 ggcttcaagg cgcacctcc ccaactgtccc tcaccccagg cccaccacgc gggcctctgc 900
 tggccactca gtcccgggat tctctgagtg ctcaaggggc ctcagggaag ccaactcactc 960
 atccattcac tcagcaaaca tttgccaagg ccgtgtcctc accaggctcg cctggcactg 1020
 ggggtataaa gaagaggcca ggtttcagct gagtgcgtat aactgtggg gggggctggg 1080

115

gcagaggcag atagtggcag ttcactttgc catctgcatg gatggggaca catgcaggtc 1140
 attgtgcgta ggcacctgcc cagaggtgag gtgagaggga tgtgcttcag ggaagtcttc 1200
 ctagaggagg caatgtctga gctaagtctt aaaggatgaa ggacaattgg ccaagtggag 1260
 acaaaggagg gaggggatta caggttgagg gctcaacatg agcaaagaca ggagatgcag 1320
 atggctgggg taatttgtca cagtagcaac aggaaatcag tacaagccct ggacctggcc 1380
 cattcttcat gtccccttcc caagcctcca agcccacatg ggcaactgcc aagatcagag 1440
 ctccaggggc ctccaggac ggggttttca gtcccttggg accatgatcc acagagagaa 1500
 attgatccta cttgagacac aggaagcaac acatgtaatg caacagccaa gcacccaacg 1560
 atcatcta at gcaacagcca aacacccagc gatcatctaa tgcaacagcc aaacacccag 1620
 cgatcatcta atgcaacagc caaacaccca gcgatcatct aatgcaacag ccaaacaccc 1680
 agcgatcatc taatgcaaca gccaaacacc cagtgatcat ctaatgcaac agccaaacac 1740
 ccagtgatca tcta atgcaa cagccaaaca gacaagtgat catcta atgc aacagcca 1798

<210> 85
 <211> 3099
 <212> DNA
 <213> Homo sapien

<400> 85
 ggcgaatggg cctctagatg catgctcgag cggcgcagtg tgatggatcg ccgcggcagg 60
 gtcacaggcg agagccctgg gatgcaccgg ccagaggcca tgctgctgct gctcacgctt 120
 gccctcctgg ggggccccac ctgggcaggg agtaagttag tgggggtctgc cctcaatctc 180
 ccctgcctcc ctccaggaga gccagggact caccgggccc ttgtcccaga ctaactctgg 240
 tcacagaacc atcctgtctg cctggagggg cgggggtcccc tgttccggca gaggtcaccc 300
 ccatatcacc gcatggggat tttcttccct ttgggtctct cttttcttca gagatgtatg 360
 gccctggagg aggcaagtat ttcagacca ctgaagacta cgaccatgaa atcacagggc 420
 tgcgggtgtc tgtaggtctt ctccctggta aaaggtagt agggctatgg tcatgggccc 480
 agcgccatgt cccctcccat cccacagttt caggaactca gggcagcggg taagcaccg 540
 tggccacttt tgccacacat gcctggctac tgtcgatgct tcctggctcc cgctgatgct 600
 tcctggctgg agcggacacg gtcagaccgt cctccctacc ttctcccttc aacccaagct 660
 caactcaacc aaaaatggcc cctctgtccc catgcctgat aggaaagtca ggggaaagtc 720
 tgtccgatta ctgtcaaaga agacaggagg taagggtcag agtggaccac tgactgaata 780
 tgagtcgcag aagtgttaga ggcagaagtc cagggccatt tccttaatat cgaagtgtct 840
 ctgctggagg tctgggatgg atttttgccc tgcatttaga agttctgggg tcctgggaga 900

ggggagagaa gcccaatagc agaggagaca gagtgtgggc ggggagagcc ggaggggtgc 960
atcctgggag agcaccaggg tgagggaggg gtgaagatga gccccgtcag ggaagcgctg 1020
gcgagtgtgg gaagtcacct gcccctcggc ctgtgagctg ctctgcttgg agtgactaag 1080
gctcgggagg tccaggctcg gccagaggca gctcatatgt gggccacagt gacggcagct 1140
gggtgccttct gggtcacgga gacctggcgc tgcacgcagc tctcctcacc aggatctcag 1200
tgactcctcc caaaagtcac acccactttg cagacgggga aactgagtcc ggagaggctg 1260
ggtaacgagc tcaagatcac agggcccaaa agtggtagaa tcagggttgg tgaccagtga 1320
gtctgtgtca gagacccaaa gtctgatggg gctggactct ctgcatcccg ggaaggagga 1380
tgggggcgct gaggaccggt gatgtgctgg gccatcccag atctggacgt ccaaagcttt 1440
gcctctctcc cagtgtccag gtgaaacttg gagactcctg ggacgtgaaa ctgggagcct 1500
taggtgggaa taccaggaa gtcaccctgc agccaggcga atacatcaca aaagtctttg 1560
tcgccttcca agctttctc cggggtatgg tcatgtacac cagcaaggac cgctatttct 1620
attttgggaa gcttgatggc cagatctcct ctgcctacc cagccaagag gggcaggtgc 1680
tggtgggcat ctatggccag tatcaactcc ttggcatcaa gagcattggc tttgaatgga 1740
attatccact agaggagccg accactgagc caccagttaa tctcacatac tcagcaaact 1800
caccctggg tgcctagggt ggggtatggg gccatccgag ctgaggccat ctgggtgggtg 1860
gtggctgatg gtactggagt aactgagtcg ggacgctgaa tctgaatcca ccaataaata 1920
aagcttctgc agaatcagt catccaggat tggctccttg atctggggta caaccaagc 1980
cttccctgct ccttgagagc aaagtcccca gtgctgcagc ccagtgaact gagatgaggg 2040
gtagggcaaa ggtgactctg ccgaggacag aaagagagca gcaccacccc ctgagagggtg 2100
ctgtggatct ctgtgccagt cccacaatct ttgaagagtc aggcttcaag gccgccactc 2160
cccactgtcc ctaccccag gccaccccag cgggcctctg ctggccactc agtcccggga 2220
ttctctgagt gctcaagggg cctcagggaa gccactcact catccattca ctgagcaaac 2280
atttgccaag gccgtgtcct caccaggctc gcctggcact ggggggtataa agaagaggcc 2340
aggtttcagc tgagtgcgta tacactgtgg tgggggctgg tgcagaggca gatagtggca 2400
gttcactttg ccatctgcat ggatggggac acatgcaggc cattgtgcgt aggcacctgc 2460
ccagagggtga ggtgagaggg atgtgcttca gggaagtctt cctagaggag gcaatgtctg 2520
agctaagtct taaaggatga aggacaattg gccaaagtga gacaaaggga ggaggggatt 2580
acaggttgag ggctcaacat gagcaaagac aggagatgca gatggctggg gtaatttgc 2640
acagtagcaa caggaaatca gtacaagccc tggacctggc ccattcttca tgtccccctc 2700

117

ccaagcctcc aagcccatat gggcacttgc caagatcaga gctccagggg cctccagggg 2760
 cgggggttttc agtccttttg gaccatgata cacagagaga aattgatcct acttgagaca 2820
 caggaagcaa cacatgtaat gcaacagcca agcacccaac gatcatctaa tgcaacagcc 2880
 aaacaccag cgatcatcta atgcaacagc caaacacca gcgatcatct aatgcaacag 2940
 ccaaaccacc agcgatcata taatgcaaca gccaaacacc cagcgatcat ctaatgcaac 3000
 agccaaacac ccagtgatca tctaatagaa cagccaaaca cccagtgate atctaatagca 3060
 acagccaaac agacaagtga tcatctaata caacagcca 3099

<210> 86
 <211> 1547
 <212> DNA
 <213> Homo sapien

<400> 86
 cggcttagcg tggcgcggc gagtcctccg gggatatgtc atgtacacca gcaaggaccg 60
 ctatttctat tttgggaagc ttgatggcca gatctcctct gcctacccca gccaaagagg 120
 gcaggtgctg gtgggcatct atggccagta tcaactcctt ggcatcaaga gcattggctt 180
 tgaatggaat tatccactag aggagccgac cactgagcca ccagttaata tcacatactc 240
 agcaaactca cccgtgggtc gctaggggtg ggtatggggc catccgagct gaggccatct 300
 ggggtgggtg ggctgatggc actggagtaa ctgagtcggg acgctgaata tgaatccacc 360
 aataaataaa gcttctgcag aatcagtgca tccaggattg gtccttggat ctgggggtaca 420
 accaaagcct tccctgctcc ttggagacaa agtccccagt gctgcagccc agtgaactga 480
 gatgaggggt agggcaaagg tgactctgcc gaggacagaa agagagcagc accacccctt 540
 cagaggtgct gtggatctct gtgccagtcc cacaatcttt gaagagtcag gcttcaaggc 600
 cgccactccc cactgtccct caccacaggc ccaccacagc ggcctctgct ggccactcag 660
 tcccgggatt ctctgagtgc tcaaggggccc tcagggaagc cactcactca tccattcact 720
 cagcaaacat ttgccaaggc cgtgtcctca ccaggctcgc ctggcactgg ggggtataaa 780
 aagaggccag gtttcagctg agtgcgtata cactgtgtgt ggggctggtg cagaggcaga 840
 tagtggcagt tcactttgcc atctgcatgg atggggacac atgcaggtea ttgtgcgtag 900
 gcacctgccc agaggtgagg tgagagggat gtgcttcagg gaagtcttcc tagaggaggc 960
 aatgtctgag ctaagtctta aaggatgaag gacaattggc caagtggaga caaaggaggg 1020
 aggggattac aggttgaggc ctcaacatga gcaaagacag gagatgcaga tggctggggg 1080
 aatttgcac agtagcaaca ggaaatcagt acaagccctg gacctggccc attcttcatg 1140
 tccccttccc aagcctccaa gccacatgg gcacttgcca agatcagagc tccagggggc 1200

118

tccagggacg gggttttcag tcctttggga ccatgatcca cagagagaaa ttgatcctac 1260
 ttgagacaca ggaagcaaca catgtaatgc aacagccaag cacccaacga tcatcctaag 1320
 caacagccaa acacccagcg atcatccta atgcaacagcca aacacccagc gatcatctaa 1380
 tgcaacagcc aaacacccag cgatcatcta atgcaacagc caaacaccca gcgatcatct 1440
 aatgcaacag ccaaacaccc agtgatcatc taatgcaaca gccaaacacc cagtgatcat 1500
 ctaatgcaac agccaaacag acaagtgatc atctaatagca acagcca 1547

<210> 87
 <211> 2345
 <212> DNA
 <213> Homo sapien

<400> 87
 catgctcgag cggcgcagtg tgatggatcg agcggccgcc cgggcaggta cgtgctccgg 60
 gatcttcagc acccgcggcc gccatcgccg tcgcttggt tcttctggac tcatctgcgc 120
 cacttgccg cttcacactc cgccgccatc atggtgaagc tcgcgaaggc aggtaaaaat 180
 caaggtgacc ccaagaaaat ggctcctcct ccaaaggagg tagaagaaga tagtgaagat 240
 gaggaatgt cagaagatga agaagatgat agcagtggag aagaggctgt catacctcag 300
 aagaaaggca agaaggctgc tgcaacctca gcaaagaagg tggctgtttc cccaacaaaa 360
 aaggttgcag ttgccacacc agccaagaaa gcagctgtca ctccaggcaa aaaggcagca 420
 gcaacacctg ccaagaagac agttacacca gccaaagcag ttaccacacc tggcaagaag 480
 ggagccacac caggcaaagc attggtagca actcctggta agaaggggtgc tgccatccca 540
 gccaaagggg caaagaatgg caagaatgcc aagaaggag acagtgatga agaggaggat 600
 gatgacagtg aggaggatga ggaggatgac gaggacgagg atgaggatga agatgaaatt 660
 gaaccagcag cgatgaaagc agcagctgct gcccctgcct cagaggatga ggacgatgag 720
 gatgacgaag atgatgagga tgacgatgac gatgaggaag atgactctga agaagaagct 780
 atggagacta caccagccaa aggaaagaaa gctgcaaaaag ttgttcctgt gaaagccaag 840
 aacgtggctg aggatgaaga tgaagaagag gatgatgagg acgaggatga cgacgacgac 900
 gaagatgatg aagatgatga tgatgaagat gatgaggagg aggaagaaga ggaggaggaa 960
 gagcctgtca aagaagcacc tggaaaacga aagaaggaaa tggccaaaca gaaagcagct 1020
 cctgaagcca agaaacagaa agtggaaggc acagaaccga ctacggcttt caatctcttt 1080
 gttggaacc taaactttaa caaatctgct cctgaattaa aaactgggtat cagcgatgtt 1140
 ttgtctaaaa atgatcttg tgttggtgat gtcagaattg gtatgactag gaaatttggt 1200
 tatgtggatt ttgaatctgc tgaagacctg gagaaagcgt tggaactcac tggtttga 1260

119

gtctttggca atgaaattaa actagagaaa ccaaaaggaa aagacagtaa gaaagagcga 1320
 gatgcgagaa cacttttggc taaaaatctc cttacaaag tcactcagga tgaattgaaa 1380
 gaagtgtttg aagatgctgc ggagatcaga ttagtcagca aggatgggaa aagtaaaggg 1440
 attgcttata ttgaatttaa gacagaagct gatgcagaga aaacctttga agaaaagcag 1500
 ggaacagaga tcgatgggcg atctatttcc ctgtactata ctggagagaa aggtcaaaat 1560
 caagactata gaggtggaaa gaatagcact tggagtgggtg aatcaaaaac tctggtttta 1620
 agcaacctct cctacagtgc aacagaagaa actcttcagg aagtatttga gaaagcaact 1680
 tttatcaaag taccacagaa ccaaaatggc aaatctaaag ggtatgcatt tatagagttt 1740
 gcttcattcg aagacgctaa agaagcttta aattcctgta ataaaaggga aattgagggc 1800
 agagcaatca ggctggaggc gaggaggctt ccgaggaggc agaggaggag gaggtgacca 1860
 caagccacaa ggaaagaaga cgaagtttga atagcttctg tccctctgct tcccttttcc 1920
 catttgaaag aaaggactct ggggttttta ctgttacctg atcaatgaca gagccttctg 1980
 aggacattcc aagacagtat acagtcctgt ggtctccttg gaaatccgtc tagttaacat 2040
 ttcaagggca ataccgtgtt ggttttgact ggatattcat ataaactttt taaagagttg 2100
 agtgatagag ctaaccctta tctgtaagtt ttgaatttat attgtttcat cccatgtaca 2160
 aaaccatttt ttctacaaa tagtttgggt ttgttggtg tttctttttt ttgttttggt 2220
 ttgttttttt ttttttttgc gttcgtgggg ttgtaaaaga aaagaaagca gaatgtttta 2280
 tcatggtttt tgcttcagcg gctttaggac aaattaaaag tcaactctgg tgccaaaaaa 2340
 aaaaaa 2345

<210> 88
 <211> 1716
 <212> DNA
 <213> Homo sapien

<400> 88
 catgctcgag cggcgcagtg tgatggatcg agcggccgcc cgggcaggta cgtgctccgg 60
 gatcttcagc acccgcggcc gccatcgccg tcgcttggtt tcttctggac tcactcgccg 120
 cacttgctcg cttcacactc cgccgccatc atggtgaagc tcgcgaaggc aggtaaaaat 180
 caaggtgacc ccaagaaaat ggctcctcct ccaaaggagg tagaagaaga tagtgaagat 240
 gaggaaatgt cagaagatga agaagatgat agcagtggag aagaggtcgt catacctcag 300
 aagaaaggca agaaggctgc tgcaacctca gcaaagaagg tggctggttc cccaacaaaa 360
 aagggttcag ttgccacacc agccaagaaa gcagctgtca ctccaggcaa aaaggcagca 420
 gcaacacctg ccaagaagac agttacacca gccaaagcag ttaccacacc tggcaagaag 480

120

ggagccacac caggcaaagc attggtagca actcctggta agaaggggtgc tgccatccca 540
 gccaaagggg caaagaatgg caagaatgcc aagaaggaa acagtgatga agaggaggat 600
 gatgacagtg aggaggatga ggaggatgac gaggacgagg atgaggatga agatgaaatt 660
 gaaccagcag cgatgaaagc agcagctgct gcccctgcct cagaggatga ggacgatgag 720
 gatgacgaag atgatgagga tgacgatgac gatgaggaa atgactctga agaagaagct 780
 atggagacta caccagccaa aggaagaaa gctgcaaaag ttgttcctgt gaaagccaag 840
 aacgtggctg aggatgaaga tgaagaagag gatgatgagg acgaggatga cgacgacgac 900
 gaagatgatg aagatgatga tgatgaagat gatgaggagg aggaagaaga ggaggaggaa 960
 gagcctgtca aagaagcacc tggaaaacga aagaaggaaa tggccaaaca gaaagcagct 1020
 cctgaagcca agaaacagaa agtggaaagg acagaaccga ctacggcttt caatctcttt 1080
 gttggaacc taaactttaa caaatctgct cctgaattaa aaactggtat cagcgatggt 1140
 tttgctaaaa atgatcttgc tgttgtggat gtcagaattg gtatgactag gaaatttgg 1200
 tatgtggatt ttgaatctgc tgaagacctg gagaaagcgt tggaaactcac tggtttgaaa 1260
 gtctttggca atgaaattaa actagagaaa caaaaggaa aagacagtaa gaaagagcga 1320
 gatgcgagaa cacttttggc taaaaatctc ccttaciaaag tcactcagga tgaattgaaa 1380
 gaagtgtttg aagatgctgc ggagatcaga ttagtcagca aggatgggaa aagtaaaggg 1440
 attgcttata ttgaatttaa gacagaagct gatgcagaga aaacctttga agaaaagcag 1500
 ggaacagaga tcgatgggag atctatcttc ctgtactata ctggagagaa aggtcaaaat 1560
 caagactata gaggtggaaa gaatagcact tggagtgggt aatcaaaaac tctggtttta 1620
 agcaacctct cctacagtgc aacagaagaa actcttcagg aagtatttga gaaagcaact 1680
 tttatcaaag tacctcggcc gcgaccacgc taagcc 1716

<210> 89
 <211> 2068
 <212> DNA
 <213> Homo sapien

<400> 89
 gcagattgca gggcccgggc tgacgggaag tgggtgggag ctgcctgcac acgcggtgcc 60
 gcggggcgagg agtagaggcg gagggagggg acacgggctc attgcggtgt gcgccctgca 120
 ctctgtccct cactcgccgc cgacgacctg tctcgccgag cgcacgcctt gccgccgccc 180
 cgcagaaatg cttcggttac ccacagtctt tcgccagatg agaccggtgt ccagggtact 240
 ggctcctcat ctactcggg cttatgccaa agatgtaaaa ttgggtgcag atgcccagac 300
 cctaagtctt caaggtgtag accttttagc cgatgctgtg gccgttacia tggggccaaa 360

121

gggaagaaca gtgattattg agcagagttg gggaagtccc aaagtaacaa aagatggtgt 420
gactgttgca aagtcaattg acttaaaaga taaatacaag aacattggag ctaaacttgt 480
tcaagatgtt gccaataaca caaatgaaga agctggggat ggcactacca ctgctactgt 540
actggcacgc tctatagcca aggaaggctt cgagaagatt agcaaagggtg ctaatccagt 600
ggaaatcagg agaggtgtga tgtagctgt tgatgctgta attgctgaac ttaaaaagca 660
gtctaaacct gtgaccaccc ctgaagaaat tgcacagggt gctacgattt ctgcaaacgg 720
agacaaagaa attggcaata tcatctctga tgcaatgaaa aaagttggaa gaaaggggtgt 780
catcacagta aaggatggaa aaacactgaa tgatgaatta gaaattattg aaggcatgaa 840
gtttgatcga ggctatatatt ctccatactt tattaataca tcaaaagggtc agaaatgtga 900
attccaggat gcctatgttc tgttgagtga aaagaaaatt tctagtatcc agtccattgt 960
acctgctctt gaaattgcc aatgctcaccg taagcctttg gtcataatcg ctgaagatgt 1020
tgatggagaa gctctaagta cactcgtctt gaataggcta aaggttgggtc ttcagggtgt 1080
ggcagtcaag gctccagggt ttggtgacaa tagaaagaac cagcttaaag atatggctat 1140
tgctactggt ggtgcagtgt ttggagaaga gggattgacc ctgaatcttg aagacgttca 1200
gcctcatgac ttaggaaaag ttggagaggt cattgtgacc aaagacgatg ccatgctctt 1260
aaaaggaaaa ggtgacaagg ctcaaattga aaaacgtatt caagaaatca ttgagcagtt 1320
agatgtcaca actagtgaat atgaaaagga aaaactgaat gaacggcttg caaaactttc 1380
agatggagtg gctgtgctga aggttggtgg gacaagtgat gttgaagtga atgaaaagaa 1440
agacagagtt acagatgccc ttaatgtac aagagctgct gttgaagaag gcattgtttt 1500
gggaggggggt tgtgccctcc ttcgatgcat tccagccttg gactcattga ctccagctaa 1560
tgaagatcaa aaaattggta tagaaattat taaaagaaca ctcaaaattc cagcaatgac 1620
cattgctaag aatgcagggt ttgaaggatc tttgatagtt gagaaaatta tgcaaagttc 1680
ctcagaagtt ggttatgatg ctatggctgg agattttgtg aatatggtgg aaaaaggaat 1740
cattgaccca acaaaggtaa atggttgaac actttctaaa aattcttttg ttgttcattt 1800
tgttaatgta tcctcacatc ccaaatcttt gatattgtca ccagaaaaat agtctaatag 1860
cagtagcaac agcaatagct cttgtgcccc tcaacaattc tcatgtacat ttccagcacc 1920
tagatatatta atactccact aaaaggaagc aggatccctt gagtagcagt tgagtccatg 1980
cctggaggca ggaaaaaatc aggatggatc tggaacaccc tgttcctgcc tggcagtaag 2040
aatattttcc aagatacgca gaacagggt 2068

<210> 90

<211> 366

122

<212> DNA
 <213> Homo sapien

<400> 90
 gccagtgtga tggattcgcg gcgagggtgc tactgttcac gaaatcctgt gcaagctcag 60
 cttggagggt gatcactcta ccccccaag tgcataatggg tctgtcaaag cctataactaa 120
 ctttgatgct gagcgggatg ctttgaacat tgaaacagcc atcaagacca aagagggtgtg 180
 gatgagggtca ccattgtcaa ctttttgacc aaccgcagca atgcacagag acaggatatt 240
 gccttcgcct accagagaag gacaaaaaag gaacttgcac cagcactgaa gtcagcctta 300
 tctggccacc tggagacggt gattttgggc ctattgaaga cacctgctca gtatgacgct 360
 tctgag 366

<210> 91
 <211> 1346
 <212> DNA
 <213> Homo sapien

<400> 91
 catgctcgag cggcgagtg tgatggatcg tggtcgaggc cgaggagcgc tctcagctct 60
 cggcgacagg ccagcttcc ttcaaatgt ctactgttca cgaaatcctg tgcaagctca 120
 gcttgagggt tgatcactct acacccccaa gtgcataatg gtctgtcaaa gcctatacta 180
 actttgatgc tgagcgggat gctttgaaca ttgaaacagc catcaagacc aaagggtgtg 240
 atgagggtcac cattgtcaac attttgacca accgcagcaa tgcacagaga caggatattg 300
 ccttcgccta ccagagaagg accaaaaagg aacttgcac agcactgaag tcagccttat 360
 ctggccacct ggagacggtg attttgggcc tattgaagac acctgctcag tatgacgctt 420
 ctgagctctg ctccagaacc aaccaggagc tgcaggaaat taacagagtc tacaaggaaa 480
 tgtacaagac tgatctggag aaggacatta tttcggacac atctggtgac ttccgcaagc 540
 tgatggttgc cctggcaaag ggtagaagag cagaggatgg ctctgtcatt gattatgaac 600
 tgattgacca agatgctcgg gatctctatg acgctggagt gaagaggaaa ggaactgatg 660
 ttcccaagtg gatcagcatc atgaccgagc ggagcgtgcc ccacctccag aaagtatttg 720
 ataggtacaa gagttacagc ccttatgaca tgttggaag catcaggaaa gaggttaaag 780
 gagacctgga aaatgctttc ctgaacctgg ttcagtgcac tcagaacaag cccctgtatt 840
 ttgctgatcg gctgtatgac tccatgaagg gcaaggggac gcgagataag gtctgatca 900
 gaatcatggt ctcccgcagt gaagtggaca tgttgaaaat taggtctgaa ttcaagagaa 960
 agtacggcaa gtccctgtac tattatatcc agcaagacac taagggcgac taccagaaag 1020
 cgctgctgta cctgtgtggt ggagatgact gaagcccgac acggcctgag cgtccagaaa 1080

123

tgggtgctcac catgcttcca gctaacaggt ctagaaaacc agcttgcgaa taacagtccc 1140
 cgtggccatc cctgtgaggg tgacgttagc attaccccca acctcatttt agttgcctaa 1200
 gcattgcctg gccttctctg ctagtctctc ctgtaagcca aagaaatgaa cattccaagg 1260
 agttggaagt gaagtctatg atgtgaaaca ctttgcctcc tgtgtactgt gtcataaaca 1320
 gatgaataaa ctgaatttgt acttta 1346

<210> 92
 <211> 756
 <212> DNA
 <213> Homo sapien

<400> 92
 ccgaggcccc tccttgctcg acgccgatat ctctgccggg tgactagctg cttcctttct 60
 ctctcgcgcg cgggtgtggtg gcagcaggcg cagcaggcgc acccagcctc gaaatgcaga 120
 acgacgccgg cgagttcgtg gacctgtacg tgccgcggaa atgctccgct agcaatcgca 180
 tcatcgggtg caaggaccac gcatccatcc agatgaacgt ggccgaggtt gacaagggtca 240
 caggcaggtt taatggccag tttaaaactt atgctatctg cggggccatt cgtaggatgg 300
 gtgagtcaga tgattccatt ctccgattgg ccaaggccga tggcatcgtc tcaaagtaag 360
 gttgggggct cacatttggg cagagtgagt ggactaggac tgctccagag gcgtgggtctt 420
 aacgttgtcc ttttccctg gttctaggaa cttttgactg gagagaatca cagatgtgga 480
 atatttgtca taaataaata atgaaaacct aaaaaaaaaa aaaaaaaac tcgagactag 540
 cttctctcaa ataataacca tacacaacac taaggggcga acctgatctc ttatacaagt 600
 atccttagtc atttcttttg tgcgcacaat taacctctc ggactccggc tcaactcattt 660
 acaccaacca cccaatatct ttaaacctag catgggcac cttttatgag gggggcggga 720
 taaaggagtc ggcctaagat aatatggcct agccat 756

<210> 93
 <211> 1420
 <212> DNA
 <213> Homo sapien

<400> 93
 ccgaggcccc tccttgctcg acgccgatat ctctgccggg tgactagctg cttcctttct 60
 ctctcgcgcg cgggtgtggtg gcagcaggcg cagcaggcgc acccagcctc gaaatgcaga 120
 acgacgccgg cgagttcgtg gacctgtacg tgccgcggaa atgctccgct agcaatcgca 180
 tcatcgggtg caaggaccac gcatccatcc agatgaacgt ggccgaggtt gacaagggtca 240
 caggcaggtt taatggccag tttaaaactt atgctatctg cggggccatt cgtaggatgg 300
 tgagtgtttc cctgggcttt gtcatcact tcgggacatc gtggacttta ccgtgcgcat 360

124

tggagtgtgt gatggtgcct gagtagatct gctggcagag tagtttgagc cagctggact 420
 gggctggccg cctgccgctt cttgaggggtg gaagaggggt gctctgagaa gacactcagg 480
 cagcagactc tgcctctcac taggaggtgc ccccccgacc ccgctccacc atagtcaggc 540
 tgcaggctgc cccgggagag gtggctcccc ttctgcgcct gtctccattc gctcagcggg 600
 ggagagacgt gggctggtgg cacagctgac cttctgccat ctcaggcagc cggagtggaa 660
 atattcttag tgtgcttttt ttttttctt aaggggtgagt cagatgattc cattctccga 720
 ttggccaagg ccgatggcat cgtctcaaag taagggtggg ggctcacatt tgggcagagt 780
 gagtggacta ggactgctcc agaggcgtgg tcttaacgtt gtccttttcc cctggttcta 840
 ggaacttttg actggagaga atcacagatg tggaatattt gtcataaata aataatgaaa 900
 acctaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 960
 aaaaaaaaaa aaaagagggg ggggcgcgcc caaaaaatcc ccccgggggg cgcgctttg 1020
 cgccccgct tttgtgtgaa aggggggccc ccatgagggg ctttttaaag ggccgcgcag 1080
 cggggcgcgc gttataaaaa gcaccagcag cgagtggggg aacacagccg agccacgcgg 1140
 gggagatctc ttggggagaa ggagcccca tatcggcggg gggggggaga gcaaaattag 1200
 ggggggaaac accctacacc taaacgaaga tatattaaga aactcttggg aggggaagta 1260
 atatataaac ttttcagaga ggggtatat aggtgggtga aaaaccagag acgcgagatc 1320
 gtatggatgt ggggtggtaa aagaatattg tgggtctagc gagtgatgt aatttcgacg 1380
 aaactttatt atagcgaggg gcgttttaga tgataagggtg 1420

<210> 94
 <211> 1536
 <212> DNA
 <213> Homo sapien

<400> 94
 ggcacgaggc atcgcgcgcg gtgtggtggc agcagggcgca gcaggcgcac ccagcctcga 60
 aatgcagaac gacgccggcg agttcgtgga cctgtacgtg ccgcggaat gctccgctag 120
 caatcgcatc atcggtgcca aggaccacgc atccatccag atgaacgtgg ccgaggtgag 180
 ctgggagccc gggaggcggg aaggttgtga tatatgtgcg ggaaaggcag gctgtcccat 240
 tgtggaggag ccctggggt gaaggtacag gcagaggctg gctttgagga ttggtgtttc 300
 ccaaacctgg gggagtgggt tgtgaccctt cttctcttcc taggttgaca aggtcacagg 360
 caggtttaat ggccagttta aaacttatgc tatctgcggg gccattcgta ggatggtgag 420
 tgtttccctg ggctttgctc atcacttcgg gacatcgtag actttaccgt gcgcattgga 480
 gtgtgtgatg gtgcctgagt agatctgctg gcagagtagt ttgagccagc tggactgggc 540

125

tggccgcctg ccgcttcttg aggggtggaag aggggtgctc tgagaagaca ctcaggcagc 600
 agactctgcc tctcactagg aggtgcccc cgcacccgc tccaccatag tcaggctgca 660
 ggctgccccg ggagaggtgg ctcccccttct gcgcctgtct ccattcgctc agcgggggag 720
 agacgtgggc tgggtggcaca gctgaccttc tgccatctca ggcagccgga gtggaaatat 780
 tcttagtgtg cttttttttt tttcttaagg gtgagtcaga tgattccatt ctccgattgg 840
 ccaaggccga tggcatcgct tcaaagtaag gttgggggct cacatttggg cagagtgagt 900
 ggactaggac tgctccagag gcgtgggtctt aacgttgctc ttttcccctg gttctaggaa 960
 cttttgactg gagagaatca cagatgtgga atatttgtca taaataaata atgaaaacct 1020
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1080
 aaaaaaaaaa gagggggggg cgcgcccaca aaatccccc ggggggcgcg cctttgcgcc 1140
 cccgcttttg tgtgaaaggg gggcccccac gaggggcttt ttaaagggcc gcgcagcggg 1200
 gcgcgcgtta taaaaagcac cagcagcgag tgggggaaca cagccgagcc acgcggggga 1260
 gatctcttgg ggagaaggag ccccatatc ggcggggggg gggagagcaa aattaggggg 1320
 ggaaacaccc tacacctaaa cgaagatata ttaagaaact cttgggaggg gaagtaatat 1380
 ataaactttt cagagagggg gtatataggt ggggtgaaaa ccagagacgc gagatcgat 1440
 ggatgtgggg tggtaaaaga atattgtggg tctagcgagt gtatgtaatt tcgacgaaac 1500
 tttattatag cgaggggcgt tttagatgat aagggtg 1536

<210> 95
 <211> 930
 <212> DNA
 <213> Homo sapien

<400> 95
 agatcatgcc gagcggcgcc agtgtgatgg atgcgtggtc gcggccgagg tacgtgccgc 60
 ggaaatgctc cgctagcaat cgcatcatcg gtgccaagga ccacgcatcc atccagatga 120
 acgtggccga ggttgacaag gtcacaggca ggtttaatgg ccagttaaa acttatgcta 180
 tctgcggggc cattcgtagg atgggtgagt cagatgatc cattctccga ttggccaagg 240
 ccgatggcat cgtctcaaag taagggtggg ggctcacatt tgggcagagt gaggggacta 300
 ggactgctcc agaggcgtgg tcttaacgtt gtccttttcc cctggttcta ggaacttttg 360
 actggagaga atcacagatg tggaatatat gtcataaata aataatgaaa acctaaaaaa 420
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 480
 aaaagagggg ggggcgcgcc caaaaaatcc ccccgggggg cgcgcctttg cgccccgct 540
 tttgtgtgaa aggggggccc ccatgagggg ctttttaaaag ggccgcgcag cggggcgcg 600

126

```

gttataaaaa gcaccagcag cgagtggggg aacacagccg agccacgcgg gggagatctc   660
ttggggagaa ggagcccca ttcggcggg gggggggaga gcaaaattag ggggggaaac   720
accctacacc taaacgaaga tatattaaga aactcttggg aggggaagta atatataaac   780
ttttcagaga ggggggtatat aggtgggtga aaaaccagag acgcgagatc gtatggatgt   840
ggggtggtaa aagaatattg tgggtctagc gagtgtatgt aatttcgacg aaactttatt   900
atagcgaggg gcgtttttaga tgataaggtg                               930

```

```

<210> 96
<211> 185
<212> PRT
<213> Homo sapien

```

```

<220>
<221> MISC_FEATURE
<222> (35)..(35)
<223> X=any amino acid

```

```

<220>
<221> MISC_FEATURE
<222> (70)..(70)
<223> X=any amino acid

```

```

<220>
<221> MISC_FEATURE
<222> (81)..(81)
<223> X=any amino acid

```

```

<220>
<221> MISC_FEATURE
<222> (93)..(93)
<223> X=any amino acid

```

```

<400> 96

```

```

Gln Lys Ser Ile His Ala Cys Asn Val Gly Gly Arg Leu Leu Cys Gln
1           5           10           15

```

```

Asp Arg Pro Pro Thr Leu Gln Lys Ser Ile His Ala Cys Ala Ala Arg
          20           25           30

```

```

Ile Ala Xaa Ser Ser Gly His Arg Pro Gly Thr Phe Ser Arg Val Thr
          35           40           45

```

```

Ala Leu Asn Asp Val Glu Thr Arg Asp Ser Thr Trp Pro His Ala Arg
          50           55           60

```

127

Cys Glu Gly Pro Ala Xaa Ser Arg Asp Val Trp Thr Pro Ala Gly Cys
65 70 75 80

Xaa Gln Glu Ala Val Glu Leu Val Gln Tyr Ala Tyr Xaa Ser Glu Lys
85 90 95

Val Arg Gly Glu Arg Arg Arg Thr Arg Lys Glu Ala Asn Val Lys Asp
100 105 110

Glu Val Lys Asp Arg Gln Ile Asp Arg Gly Glu Thr Ala Lys Arg Thr
115 120 125

Leu Glu Gln Lys Arg Lys Arg Arg Lys Thr Arg Gln Pro Asp Ala Lys
130 135 140

Asp Gly Asp Ser Tyr Asp Pro Tyr Asp Phe Ser Asp Thr Glu Glu Glu
145 150 155 160

Met Pro Gln Val His Thr Pro Lys Thr Ala Asp Ser Gln Glu Thr Lys
165 170 175

Glu Ser Gln Lys Val Glu Leu Ser Glu
180 185

<210> 97
<211> 109
<212> PRT
<213> Homo sapien

<220>
<221> MISC_FEATURE
<222> (11)..(11)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (39)..(39)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (41)..(42)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (44)..(44)
<223> X=any amino acid

128

<220>
 <221> MISC_FEATURE
 <222> (55)..(57)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (77)..(77)
 <223> X=any amino acid

<400> 97

Ala Glu Thr Cys Gly Pro Leu Gln Asp Ala Xaa Arg Lys Leu Trp Ser
 1 5 10 15

Trp Ser Ser Met Leu Thr Phe Arg Glu Gly Ser Trp Arg Thr Glu Lys
 20 25 30

Lys Arg Lys Lys Arg Ser Xaa Gly Xaa Xaa Gln Xaa Gln Lys Met Lys
 35 40 45

Arg Arg Lys Ala Lys Arg Xaa Xaa Xaa Arg Arg Gly Arg Glu Gly Arg
 50 55 60

Leu Ala Ser Gln Met Pro Lys Met Gly Ile His Thr Xaa Pro Met Thr
 65 70 75 80

Ser Val Thr Gln Arg Arg Lys Cys Leu Lys Tyr Thr Leu Gln Arg Arg
 85 90 95

Gln Thr His Arg Arg Pro Arg Asn Pro Arg Lys Trp Ser
 100 105

<210> 98
 <211> 106
 <212> PRT
 <213> Homo sapien

<400> 98

Pro Gly Leu Ile Pro Leu Glu Asp Lys Glu Asp Tyr Gly Pro Asn Lys
 1 5 10 15

Glu Cys Pro Leu Cys Leu Cys Pro Arg Leu Phe Glu Ser Leu Ser Arg
 20 25 30

Asp Leu Lys Lys Asp Tyr Gly Val Tyr Leu Glu Asp Ser Gly Thr His
 35 40 45

129

Cys Leu Glu Val Ser Val Gln Ile Phe Ile Asp Asp Lys Gly Ile Leu
 50 55 60

Arg Gln Ile Thr Leu Asn Asp Leu Pro Val Gly Arg Ser Val Asp Glu
 65 70 75 80

Thr Leu Arg Leu Val Gln Ala Phe Gln Tyr Thr Asp Lys His Gly Glu
 85 90 95

Val Cys Pro Ala Gly Trp Lys Pro Gly Lys
 100 105

<210> 99
 <211> 75
 <212> PRT
 <213> Homo sapien

<400> 99

Ile Pro Lys Ser Arg Ser Gln Lys Asp Tyr Gly Val Tyr Leu Glu Asp
 1 5 10 15

Ser Gly His Thr Leu Arg Gly Leu Phe Ile Ile Asp Asp Lys Gly Ile
 20 25 30

Leu Arg Gln Ile Thr Leu Asn Asp Leu Pro Val Gly Arg Ser Val Asp
 35 40 45

Glu Thr Leu Arg Leu Val Gln Ala Phe Gln Tyr Thr Asp Lys His Gly
 50 55 60

Glu Val Cys Pro Ala Gly Trp Lys Pro Gly Lys
 65 70 75

<210> 100
 <211> 224
 <212> PRT
 <213> Homo sapien

<400> 100

Met Ser Tyr Leu Lys Arg Leu Cys Gly Thr Phe Leu Gly Gly Pro Lys
 1 5 10 15

Pro Pro Gln Arg Val Met Phe Thr Glu Asp Leu Lys Leu Pro Ala Ser
 20 25 30

Phe Asp Ala Arg Glu Gln Trp Pro Gln Cys Pro Thr Ile Lys Glu Ile
 35 40 45

130

Arg Asp Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Gly Ala Val Glu
50 55 60

Ala Ile Ser Asp Arg Ile Cys Ile Gln His Gln Cys Ala Arg Arg Ala
65 70 75 80

Trp Arg Cys Arg Arg Arg Thr Cys Ser His Ala Val Ala Ala Cys Val
85 90 95

Gly Thr Ala Val Met Val Ala Ile Leu Leu Lys Leu Gly Thr Ser Gly
100 105 110

Gln Glu Lys Ala Trp Phe Leu Val Ala Ile Tyr Glu Ser His Val Gly
115 120 125

Cys Arg Pro Tyr Phe His Thr Leu Pro Val Ser Thr Thr Ser Lys Gly
130 135 140

Ser Arg Pro Pro Cys Thr Gly Glu Gly Asp Thr Pro Lys Cys Ser Lys
145 150 155 160

Ser Cys Glu Pro Gly Tyr Arg Pro Ser Tyr Lys Gln Asp Lys Arg Tyr
165 170 175

Gly Tyr Asn Ser Tyr Ser Val Ser Asn Ser Glu Lys Asp Ile Met Ala
180 185 190

Glu Ile Ser Lys Asn Gly Pro Trp Arg Glu Leu Ser Leu Cys Ile Gly
195 200 205

Leu Pro Gly Leu Glu Val Arg Glu Cys Thr Asn Thr Ser Pro Glu Arg
210 215 220

<210> 101

<211> 181

<212> PRT

<213> Homo sapien

<220>

<221> MISC_FEATURE

<222> (16)..(16)

<223> X=any amino acid

<220>

<221> MISC_FEATURE

<222> (19)..(20)

<223> X=any amino acid

131

<220>
<221> MISC_FEATURE
<222> (22)..(29)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (31)..(32)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (34)..(34)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (37)..(37)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (153)..(153)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (172)..(172)
<223> X=any amino acid

<400> 101

Pro Leu Arg Gln Arg Gln Pro Leu Arg Cys Ala Gln Ala Gly Leu Xaa
1 5 10 15

Ala Leu Xaa Xaa Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Xaa Xaa
20 25 30

Ser Xaa Leu Ser Xaa Ser Leu Cys Cys Leu Leu Val Leu Ala Asn Ala
35 40 45

Arg Ser Arg Pro Ser Phe His Pro Leu Ser Asp Glu Leu Val Asn Tyr
50 55 60

Val Asn Lys Arg Asn Thr Thr Trp Gln Ala Gly His Asn Phe Tyr Asn
65 70 75 80

Val Asp Met Ser Tyr Leu Lys Arg Leu Cys Gly Thr Phe Leu Gly Gly
85 90 95

132

Pro Lys Pro Pro Gln Arg Val Met Phe Thr Glu Asp Leu Lys Leu Pro
 100 105 110

Ala Ser Phe Asp Ala Arg Glu Gln Trp Pro Gln Cys Pro Thr Ile Lys
 115 120 125

Glu Ile Arg Asp Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Gly Ala
 130 135 140

Val Glu Ala Ile Ser Asp Arg Ile Xaa Ile His Thr Asn Ala His Val
 145 150 155 160

Glu Arg Gly Gly Val Gly Gly Gly Pro Ala His Xaa Leu Trp Gln His
 165 170 175

Val Trp Gly Arg Leu
 180

<210> 102
 <211> 227
 <212> PRT
 <213> Homo sapien

<220>
 <221> MISC_FEATURE
 <222> (45)..(45)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (78)..(78)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (85)..(85)
 <223> X=any amino acid

<400> 102

Thr Ser Leu His Glu Asp Asp Trp Arg Ser Arg Pro Ser Arg Gly Pro
 1 5 10 15

Ala Leu Thr Pro Ile Arg Asp Glu Glu Trp Gly Gly His Ser Pro Arg
 20 25 30

Ser Pro Arg Gly Trp Asp Gln Glu Pro Ala Arg Glu Xaa Ala Gly Gly
 35 40 45

133

Gly Trp Arg Ala Arg Arg Pro Arg Ala Arg Ser Asp Arg Arg His Trp
 50 55 60

Thr Thr Ser Pro Arg Arg Ala Pro His Glu Ser Gly Ser Xaa Ser Pro
 65 70 75 80

Thr Asn Asn Gly Xaa Arg Ser Arg Ala Tyr Met Pro Thr Val Asp Pro
 85 90 95

His Val Arg Asp Asp Leu Leu Trp Thr Lys Tyr Asn Ser Arg Asp Ile
 100 105 110

Pro Thr Ala Thr Thr Gly Asp Pro Leu Leu Leu Tyr Asn Ile Gln Ala
 115 120 125

Leu Arg Asp Ala Ala Leu Leu Ser Tyr Pro Met Val Pro Thr His His
 130 135 140

Ala Tyr Leu Gly Thr Leu Trp Asp Lys Arg Leu Pro Gly Ser Gly Asp
 145 150 155 160

Leu Pro Tyr Asp Gly Arg Leu Leu Glu Glu Ala Val Arg Lys Lys Gly
 165 170 175

Gly Arg Arg Arg Arg Arg Ile Pro His Lys Glu Glu Glu Glu Glu Ala
 180 185 190

Tyr Tyr Pro Pro Ala Pro Pro Pro Tyr Ser Glu Thr Asp Ser Gln Ala
 195 200 205

Ser Arg Glu Arg Arg Leu Lys Lys Asn Leu Ala Leu Ser Arg Glu Ser
 210 215 220

Leu Val Val
 225

<210> 103
 <211> 222
 <212> PRT
 <213> Homo sapien

<220>
 <221> MISC_FEATURE
 <222> (10)..(10)
 <223> X=any amino acid

134

<220>
 <221> MISC_FEATURE
 <222> (45)..(45)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (59)..(59)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (79)..(79)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (86)..(86)
 <223> X=any amino acid

<400> 103

Ser Pro Pro Ser Thr Arg Thr Ile Gly Xaa Leu Gly Leu Pro Gly Ala
 1 5 10 15

Leu Pro Ser Pro Arg Ser Gly Met Arg Ser Gly Val Ala Thr Pro Pro
 20 25 30

Gly Val Pro Gly Asp Gly Thr Arg Ser Pro Pro Gly Xaa Arg Gln Ala
 35 40 45

Gly Ala Gly Gly Pro Gly Gly Pro Gly Pro Xaa Pro Ile Asp Ala Thr
 50 55 60

Gly Arg Pro His Pro Ala Glu His Arg Thr Ser Gln Gly Ala Xaa Leu
 65 70 75 80

Pro Arg Ile Met Val Xaa Glu Ala Gly His Tyr Met Pro Pro Gln Ser
 85 90 95

Pro Ser Arg Asp Asp Leu Tyr Asp Gln Asp Asn Ser Arg Asp Ile Pro
 100 105 110

Thr Leu Pro Gln Ala Thr Pro Ile Tyr Asp Asn Ile Gln Ala Pro Arg
 115 120 125

Glu Arg Pro Pro Ala Tyr Pro Arg Ser His His His Arg Thr Arg Asp
 130 135 140

135

Pro Arg Asp Asn Gly Ser Arg Ser Gly Asp Leu Pro Tyr Asp Gly Arg
 145 150 155 160

Leu Leu Glu Glu Ala Val Arg Lys Lys Gly Val Gly Gly Glu Glu Asp
 165 170 175

Thr Pro Gln Gly Gly Gly Gly Arg Gly Leu Leu Pro Ala Arg Ala Ala
 180 185 190

Pro Val Leu Gly Asp Arg Leu Ala Gly Val Pro Arg Ala Gln Ala Gln
 195 200 205

Glu Glu Leu Gly Pro Glu Ser Gly Lys Phe Ser Arg Leu Ile
 210 215 220

<210> 104
 <211> 74
 <212> PRT
 <213> Homo sapien

<400> 104

Met Arg Leu Gly Val Phe Val Arg Arg Leu Leu Cys Val Pro Gly Arg
 1 5 10 15

Gly Asp Asp Val Val Leu Val Val Val Cys Leu Trp Glu Pro His Val
 20 25 30

Gly Thr Ala Val Gly Lys Tyr Tyr Arg Arg Ala Lys Cys Gly Gly Pro
 35 40 45

Ser Ser Leu Asp Gly Ile Cys Met Met Ser Ser Glu Gly Arg Asp Val
 50 55 60

Cys Gly Gly Leu Arg Phe Leu Ser Cys Ile
 65 70

<210> 105
 <211> 85
 <212> PRT
 <213> Homo sapien

<400> 105

Gly Val Cys Ser Gly Val Leu Leu Ala Trp Ser Asp Ala Ser Trp Ser
 1 5 10 15

Phe Arg Glu Ala Pro Leu Cys Val Pro Gly Arg Gly Asp Asp Val Val
 20 25 30

136

Leu Val Val Val Cys Leu Trp Glu Pro His Val Gly Thr Ala Val Gly
 35 40 45

Lys Tyr Tyr Arg Arg Ala Lys Cys Gly Gly Pro Ser Ser Leu Asp Gly
 50 55 60

Ile Cys Met Met Ser Ser Glu Gly Arg Asp Val Cys Gly Gly Leu Arg
 65 70 75 80

Phe Leu Ser Cys Ile
 85

<210> 106
 <211> 85
 <212> PRT
 <213> Homo sapien

<400> 106

Gly Val Cys Ser Gly Val Leu Leu Ala Trp Ser Asp Ala Ser Trp Ser
 1 5 10 15

Phe Arg Glu Ala Pro Leu Cys Val Pro Gly Arg Gly Asp Asp Val Val
 20 25 30

Leu Val Val Val Cys Leu Trp Glu Pro His Val Gly Thr Ala Val Gly
 35 40 45

Lys Tyr Tyr Arg Arg Ala Lys Cys Gly Gly Pro Ser Ser Leu Asp Gly
 50 55 60

Ile Cys Met Met Ser Ser Glu Gly Arg Asp Val Cys Gly Gly Leu Arg
 65 70 75 80

Phe Leu Ser Cys Ile
 85

<210> 107
 <211> 66
 <212> PRT
 <213> Homo sapien

<400> 107

Thr Gly Arg Leu Tyr Ser Pro Pro Glu Cys Arg Gly Lys Ser Leu Thr
 1 5 10 15

Ser Lys Gly Pro Thr Lys Gln Phe Arg Asn Leu Pro Pro Val Asn Val

137

20

25

30

Pro Thr Thr Glu Val Ser Pro Thr Phe Ser Glu Asn His His Lys Asn
 35 40 45

His His Thr Lys Cys Ser Ser Tyr Thr Glu Tyr Thr Cys Gln Gly Ser
 50 55 60

Ser Arg
 65

<210> 108
 <211> 66
 <212> PRT
 <213> Homo sapien
 <220>
 <221> MISC_FEATURE
 <222> (54)..(54)
 <223> X=any amino acid

<400> 108

Thr Gly Arg Leu Tyr Ser Pro Pro Glu Cys Arg Gly Lys Ser Leu Thr
 1 5 10 15

Ser Lys Gly Pro Thr Lys Gln Phe Arg Asn Leu Pro Pro Val Asn Val
 20 25 30

Pro Thr Thr Glu Val Ser Pro Thr Ser Gln Lys Thr Thr Thr Lys Thr
 35 40 45

Thr Thr Pro Asn Ala Xaa Ala Thr Arg Ser Thr Pro Ala Arg Asp Pro
 50 55 60

Leu Glu
 65

<210> 109
 <211> 126
 <212> PRT
 <213> Homo sapien

<400> 109

Met Trp His Leu Ser Pro Phe Ala Leu Gly Ile Cys Asp Pro Ser Ile
 1 5 10 15

Val Leu Arg Pro Leu Cys Pro His Phe Pro Val His Val Gly Asp Asp
 20 25 30

138

Gly Ser Pro Phe Pro Phe Ala Gln Leu Pro Pro Gly Ala Arg Gly Pro
 35 40 45

Ser Pro Gln Gly Val Trp Ile Tyr Ser Phe Ile Arg Pro Gly Pro Pro
 50 55 60

Met Phe Ala Cys Leu Cys Thr Ser Thr Pro Asn Val Ser Ala Leu Pro
 65 70 75 80

Pro Glu Ala Leu Cys Arg Ala Ser Leu Phe Trp Arg Gly Arg Gly Cys
 85 90 95

Gly Val Thr Cys Thr Leu Gly Leu Val Asp Thr Val Asn Ser Ser Gln
 100 105 110

Val Asp Phe Ser Gly Gly Glu Lys Lys Gly His Leu Arg Leu
 115 120 125

<210> 110
 <211> 117
 <212> PRT
 <213> Homo sapien

<400> 110

Leu Gly Pro Val Phe Ser Arg Ala Pro Phe Leu Thr Leu Val Trp Ile
 1 5 10 15

Thr Cys Val Gly Met Trp His Leu Ser Pro Phe Ala Leu Gly Ile Cys
 20 25 30

Asp Pro Ser Ile Val Leu Arg Pro Leu Cys Pro His Phe Pro Val His
 35 40 45

Val Gly Asp Asp Gly Ser Pro Phe Pro Phe Ala Gln Leu Pro Pro Gly
 50 55 60

Ala Arg Gly Pro Ser Pro Gln Gly Val Trp Ile Tyr Ser Phe Ile Arg
 65 70 75 80

Pro Gly Pro Pro Met Phe Ala Cys Leu Cys Thr Ser Thr Pro Asn Val
 85 90 95

Ser Ala Leu Pro Pro Glu Ala Leu Cys Arg Ala Ser Leu Phe Trp Glu
 100 105 110

139

Asp Gly Gly Ala Val
115

<210> 111
<211> 170
<212> PRT
<213> Homo sapien

<400> 111

Met Tyr Phe Lys Asp Tyr Ile Gln Glu Arg Ser Asp Pro Val Glu Gln
1 5 10 15

Gly Lys Pro Val Ile Pro Ala Ala Val Leu Gly Arg Leu His Arg Lys
20 25 30

Trp Thr Tyr Ser Ala Val Ala Val Ser Pro Gly Ala Ala Ile Thr Gln
35 40 45

Ile Leu Pro Val Ile His Gln Leu Asp Trp Arg Leu Met Glu Phe Lys
50 55 60

Leu Ala Asp Pro Asp Glu Val Ala Ala Ser Gly Glu Arg Gly Leu Ala
65 70 75 80

His Asp Glu Leu Arg Glu Ala Glu Pro Gly Leu Thr Leu Leu Arg
85 90 95

Leu Glu His His Ala Gln Asp Val Gly Glu Ser Ala Thr Cys Ser Arg
100 105 110

Leu Asn Val Arg Thr Ser Glu Thr Cys Leu Gly Phe Gln Arg Pro Glu
115 120 125

Gly Thr Val Thr Arg Ile Thr Trp Ala Val Thr Thr Pro Tyr Thr Gly
130 135 140

Arg Tyr Leu Thr Phe Arg Pro Gly Thr His Pro Leu Asn Pro Ala Pro
145 150 155 160

Gln Gly Phe Val Val Pro Val Gly Cys Pro
165 170

<210> 112
<211> 225
<212> PRT
<213> Homo sapien

<220>

140

<221> MISC_FEATURE
<222> (41)..(43)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (46)..(48)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (50)..(50)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (57)..(57)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (68)..(68)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (95)..(95)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (101)..(101)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (110)..(110)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (113)..(113)
<223> X=any amino acid

<400> 112

Lys Asp Phe Asp Ser Pro Glu Asn Gly Ala Asp Ser Phe Gln Ser Ser
1 5 10 15

Asp Ser Leu Leu Gln Ser Trp Asn Ser Gln Ser Ser Leu Leu Asp Val
20 25 30

141

Gln Arg Val Pro Ser Phe Glu Ser Xaa Xaa Xaa Asp Cys Xaa Xaa Xaa
 35 40 45

Leu Xaa Leu Asn Lys Pro Thr Cys Xaa Ser Arg Ile Thr Ser Lys Arg
 50 55 60

Gly Val Thr Xaa Trp Ser Lys Ala Asn Gln Leu Tyr Leu Gln Leu Cys
 65 70 75 80

Trp Pro Ala Ser Gln Glu Val Asp Leu Phe Ser Cys Gly Ser Xaa Ser
 85 90 95

Trp Ser Cys Tyr Xaa Thr Asn Pro Ala Ser His Ser Ser Xaa Gly Leu
 100 105 110

Xaa Thr Asp Gly Ser Leu Ser Ser Pro Thr Pro Met Arg Trp Pro Ala
 115 120 125

Ser Gly Glu Arg Gly Leu Ala His Asp Glu Leu Arg Glu Ala Glu Pro
 130 135 140

Gly Leu Thr Leu Leu Leu Arg Leu Glu His His Ala Gln Asp Val Gly
 145 150 155 160

Glu Ser Ala Thr Cys Ser Arg Leu Asn Val Arg Thr Ser Glu Thr Cys
 165 170 175

Leu Gly Phe Gln Arg Pro Glu Gly Thr Val Thr Arg Ile Thr Trp Ala
 180 185 190

Val Thr Thr Pro Tyr Thr Gly Arg Tyr Leu Thr Phe Arg Pro Gly Thr
 195 200 205

His Pro Leu Asn Pro Ala Pro Gln Gly Phe Val Val Pro Val Gly Cys
 210 215 220

Pro
 225

<210> 113
 <211> 175
 <212> PRT
 <213> Homo sapien

<400> 113

142

Gly Gly Glu Glu Gly Arg Ala Ser Trp Gly Gln Cys Arg Leu Phe Gly
 1 5 10 15

Pro Gly Lys Leu Arg Trp Ala Gly Leu Pro Pro Val Trp Leu Cys Gln
 20 25 30

Gly His Pro Gly Val Leu His Leu Gly Pro Gly Gly Trp Glu Gly Arg
 35 40 45

Glu Ala Phe Gly Leu Leu Asn His Leu Glu Val Ser Leu Leu Gln Thr
 50 55 60

Ser Ala Gly Ser Gly Ser Pro Gly Val Met Gly Ser Gly Trp Leu Asn
 65 70 75 80

Leu Glu Ile Val Trp Ser Leu Phe Glu Gly Pro Ala Trp Leu Leu Leu
 85 90 95

Gln Arg Asn Cys Arg His Leu Ser Phe Pro Ser Leu Pro His Pro Thr
 100 105 110

Ala Glu Lys Gly Trp Arg Gly Glu Ser Ser Ser Ala Phe His Ser Val
 115 120 125

Tyr Val Ser Gly Asp Ser Arg Gly Ala Gly Leu Lys Ile Ala Gly Gly
 130 135 140

Arg Pro Ser Pro Gly Cys Cys Ser Val Gly Ala Trp Pro Ser Ser Ser
 145 150 155 160

Arg Pro Thr Cys Phe Leu Trp Cys Gly Gln Ser Gln Leu Pro Ser
 165 170 175

<210> 114
 <211> 270
 <212> PRT
 <213> Homo sapien

<400> 114

Met Asp Asp Gln Arg Asp Leu Ile Ser Asn Asn Glu Gln Leu Pro Met
 1 5 10 15

Leu Gly Arg Arg Pro Gly Ala Pro Glu Ser Lys Cys Ser Arg Gly Ala
 20 25 30

Leu Tyr Thr Gly Phe Ser Ile Leu Val Thr Leu Leu Leu Ala Gly Gln
 35 40 45

143

Ala Thr Thr Ala Tyr Phe Leu Tyr Gln Gln Gln Gly Arg Leu Asp Lys
 50 55 60

Leu Thr Val Thr Ser Gln Asn Leu Gln Leu Glu Asn Leu Arg Met Lys
 65 70 75 80

Leu Pro Lys Pro Pro Lys Pro Val Ser Lys Met Arg Met Ala Thr Pro
 85 90 95

Leu Leu Met Gln Ala Leu Pro Met Gly Ala Leu Pro Gln Gly Pro Met
 100 105 110

Gln Asn Ala Thr Lys Tyr Gly Asn Met Thr Glu Asp His Val Met His
 115 120 125

Leu Leu Gln Asn Ala Asp Pro Leu Lys Val Tyr Pro Pro Leu Lys Gly
 130 135 140

Ser Phe Pro Glu Asn Leu Arg His Leu Lys Asn Thr Met Glu Thr Ile
 145 150 155 160

Asp Trp Lys Val Phe Glu Ser Trp Met His His Trp Leu Leu Phe Glu
 165 170 175

Met Ser Arg His Ser Leu Glu Gln Lys Pro Thr Asp Ala Pro Pro Lys
 180 185 190

Val Leu Thr Lys Cys Gln Glu Glu Val Ser His Ile Pro Gly Cys Pro
 195 200 205

Pro Gly Phe Ile Gln Ala Gln Val Arg Arg Glu Arg Gln Leu Ser Ala
 210 215 220

Thr Pro Val Leu Trp Gly Ala Ser Ala Thr Ala Gly Val Ser Ser Pro
 225 230 235 240

Thr Ala Arg Arg Ser Pro Thr Pro Glu Ala Ala Gly Thr Ile Thr Ala
 245 250 255

Val Ser His Trp Asn Trp Arg Thr Arg Leu Leu Gly Trp Val
 260 265 270

<210> 115
 <211> 225
 <212> PRT

144

<213> Homo sapien

<400> 115

Gly Arg Thr Gly Asp Ala Val Cys Cys Pro Pro Ala Leu Leu Asp Leu
 1 5 10 15

Arg Gly Pro Pro Gly Pro Pro Ser Ala Gly Phe Asp Phe Ser Phe Leu
 20 25 30

Pro Gln Pro Pro Gln Glu Lys Ala His Asp Gly Gly Arg Tyr Tyr Arg
 35 40 45

Ala Asp Asp Ala Asn Val Val Arg Asp Arg Asp Leu Glu Val Asp Thr
 50 55 60

Thr Leu Lys Ser Leu Ser Gln Gln Ile Glu Asn Ile Arg Ser Pro Glu
 65 70 75 80

Gly Ser Arg Lys Asn Pro Ala Arg Thr Cys Cys Asp Leu Lys Met Cys
 85 90 95

Gln Ser Asp Trp Lys Ser Gly Glu Tyr Trp Ile Asp Pro Asn Gln Gly
 100 105 110

Cys Ser Leu Asp Ala Ile Lys Val Phe Met Arg Thr Met Glu Thr Gly
 115 120 125

Glu Thr Leu Arg Val Pro His Ser Ala Ser Val Trp Arg Gln Lys Asn
 130 135 140

Trp Tyr Ile Ser Lys Asn Pro Lys Asp Lys Arg His Val Trp Phe Gly
 145 150 155 160

Glu Ser Met Thr Asp Gly Phe Gln Phe Glu Tyr Gly Gly Gln Gly Ser
 165 170 175

Asp Pro Ala Asp Val Ala Ile Gln Leu Thr Phe Leu Arg Leu Met Ser
 180 185 190

Ser Glu Ala Phe Gln Asn Ile Thr Tyr His Cys Lys Asn Ser Val Ala
 195 200 205

Tyr Met Asp His Gln Thr Gly Asn Leu Lys Lys Ala Leu Leu Leu Gln
 210 215 220

Gly

145

225

<210> 116
 <211> 121
 <212> PRT
 <213> Homo sapien

 <220>
 <221> MISC_FEATURE
 <222> (89)..(89)
 <223> X=any amino acid

<400> 116

Trp Ile Asp Pro Lys Leu Arg Leu Gln Ala Gly Cys His Pro Ile Arg
 1 5 10 15

Cys Met Arg Ser Met Glu Thr Gly Glu Thr Leu Arg Val Pro His Ser
 20 25 30

Ala Ser Val Trp Arg Gln Lys Asn Trp Tyr Ile Ser Lys Asn Pro Lys
 35 40 45

Asp Lys Arg His Val Trp Phe Gly Glu Ser Met Thr Asp Gly Phe Gln
 50 55 60

Phe Glu Tyr Gly Gly Gln Gly Ser Asp Pro Ala Asp Val Ala Ile Gln
 65 70 75 80

Leu Thr Phe Leu Arg Leu Met Ser Xaa Glu Ala Phe Gln Asn Ile Thr
 85 90 95

Tyr His Cys Lys Asn Ser Val Ala Tyr Met Asp His Gln Thr Gly Asn
 100 105 110

Leu Lys Lys Ala Leu Leu Leu Gln Gly
 115 120

<210> 117
 <211> 66
 <212> PRT
 <213> Homo sapien

<400> 117

Met Ala Leu Asn Arg Glu Phe Ser Phe Ile Val Ile Thr Thr Val Thr
 1 5 10 15

Leu Thr Leu Cys Leu Ser Gly Leu Phe Ala Ile Leu Arg Ser Val Gly
 20 25 30

146

Phe Pro Ile Phe Arg Asp Ile Ile Glu Leu Thr Ile Pro Tyr Tyr Gly
 35 40 45

Phe Ile Phe Phe Pro Phe Lys His Ser Lys Glu Thr Ile Tyr Tyr Phe
 50 55 60

Phe Pro
 65

<210> 118
 <211> 81
 <212> PRT
 <213> Homo sapien
 <220>
 <221> MISC_FEATURE
 <222> (75)..(75)
 <223> X=any amino acid

<400> 118

Gln Arg Val Leu Phe His Cys Asp His Tyr Arg Asp Thr Tyr Phe Val
 1 5 10 15

Pro Ile Arg Thr Phe Cys Asn Ile Ala Leu Cys Arg Leu Ser Asn Leu
 20 25 30

Gln Gly Tyr His Arg Ala Arg Pro Phe Pro Thr Met Asp Leu Phe Phe
 35 40 45

Phe Leu Ser Asn Thr Val Arg Lys Gln Ser Ile Thr Phe Phe Leu Lys
 50 55 60

Arg Arg Ile Tyr Ser Thr Val Ile Gln Leu Xaa Asn Ile Phe Arg Met
 65 70 75 80

Met

<210> 119
 <211> 253
 <212> PRT
 <213> Homo sapien

<400> 119

Met Val Asp Tyr Tyr Glu Val Leu Gly Val Gln Arg His Ala Leu Tyr
 1 5 10 15

147

Pro Arg Asp Ser Tyr Lys Arg His Ile Gly Lys Leu Ala Leu Lys Trp
 20 25 30

His Pro Asp Lys Asn Pro Glu Asn Lys Glu Glu Ala Ser Ser Arg Lys
 35 40 45

Phe Lys Gln Val Ala Glu Ala Tyr Glu Val Leu Ser Asp Ala Lys Lys
 50 55 60

Arg Asp Ile Tyr Asp Lys Tyr Gly Asn Arg Arg Ile Lys Val Val Glu
 65 70 75 80

Asp Gly Gly Gly Ser His Phe Asp Ser Pro Phe Glu Phe Gly Phe Thr
 85 90 95

Phe Arg Asn Pro Asp Asp Val Phe Arg Glu Phe Phe Arg Trp Lys Gly
 100 105 110

Pro Ile Leu His Leu Thr Ser Leu Lys Thr Leu Leu Arg Thr Ser Leu
 115 120 125

Gly Ile Glu Gly Val Pro Glu Glu Ala Glu Ala Glu Gly Arg Gly Arg
 130 135 140

Phe Asn Leu Arg Ser Val Asp Phe Arg Leu Leu Glu Ala Gly Trp Ser
 145 150 155 160

Ser Met Asp Ala Gly Phe Thr Ser Leu Gly Ser Leu Gly His Gly Val
 165 170 175

Leu Thr Leu Phe Ser Ser Thr Ser Phe Gly Gly Ser Gly Met Gly Asn
 180 185 190

Tyr Lys Ser Ile Ser Thr Ser Thr Lys Leu Val Asn Gly Arg Pro Ile
 195 200 205

Thr Thr Lys Arg Ile Val Asp Asn Ser Gln Asp Arg Val Gln Val Glu
 210 215 220

Asp Asp Gly Gln Leu Lys Phe Leu Thr Ile Gly Tyr Glu Gln Leu Leu
 225 230 235 240

Cys Leu Asp Asn Lys Met Ile Gln Arg Thr Arg Leu Ala
 245 250

148

<210> 120
 <211> 203
 <212> PRT
 <213> Homo sapien

 <220>
 <221> MISC_FEATURE
 <222> (32)..(32)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (104)..(104)
 <223> X=any amino acid

<400> 120

Arg Arg Ser Ser Ser Arg Lys Phe Lys Gln Val Ala Glu Ala Tyr Glu
1 5 10 15

Val Leu Ser Asp Ala Lys Lys Arg Asp Ile Tyr Asp Lys Tyr Gly Xaa
20 25 30

Arg Arg Ile Lys Trp Trp Arg Thr Glu Val Glu Val Ile Leu Thr Val
35 40 45

His Leu Asn Leu Ala Ser His Ser Val Thr Gln Met Met Ser Ser Gly
50 55 60

Asn Phe Leu Gly Gly Arg Asp Pro Phe Ser Phe Asp Phe Phe Glu Asp
65 70 75 80

Pro Phe Glu Asp Phe Phe Gly Asn Arg Arg Gly Pro Arg Gly Ser Arg
85 90 95

Ser Arg Gly Thr Gly Ser Phe Xaa Ser Ala Phe Ser Gly Phe Pro Ser
100 105 110

Phe Val Ser Gly Trp Ser Ser Met Asp Ala Gly Phe Thr Ser Leu Gly
115 120 125

Ser Leu Gly His Gly Val Leu Thr Leu Phe Ser Ser Thr Ser Phe Gly
130 135 140

Gly Ser Gly Met Gly Asn Tyr Lys Ser Ile Ser Thr Ser Thr Lys Leu
145 150 155 160

Val Asn Gly Arg Pro Ile Thr Thr Lys Arg Ile Val Asp Asn Ser Gln
165 170 175

149

Asp Arg Val Gln Val Glu Asp Asp Gly Gln Leu Lys Phe Leu Thr Ile
180 185 190

Gly Tyr Glu Gln Leu Leu Cys Leu Asp Asn Lys
195 200

<210> 121
<211> 128
<212> PRT
<213> Homo sapien

<400> 121

Met Ala Val Gln Ile Ser Lys Lys Arg Lys Phe Val Ala Asp Gly Ile
1 5 10 15

Phe Lys Ala Glu Leu Asn Glu Phe Leu Thr Arg Glu Leu Ala Glu Asp
20 25 30

Gly Tyr Ser Gly Val Glu Val Arg Val Thr Pro Thr Arg Thr Glu Ile
35 40 45

Ile Ile Leu Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg
50 55 60

Arg Ile Arg Glu Leu Thr Ala Val Val Gln Lys Arg Phe Gly Phe Pro
65 70 75 80

Glu Gly Ser Val Glu Leu Tyr Ala Glu Lys Val Ala Thr Arg Gly Leu
85 90 95

Cys Ala Ile Ser Pro Gly Arg Val Ser Ala Val Pro Thr Pro Arg Arg
100 105 110

Ala Arg Cys Ala Ala Ser Phe Leu Ser Leu Ser Arg Thr Pro Met Gly
115 120 125

<210> 122
<211> 143
<212> PRT
<213> Homo sapien

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> X=any amino acid

<220>

150

<221> MISC_FEATURE
 <222> (115)..(115)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (119)..(119)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (121)..(121)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (123)..(123)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (125)..(126)
 <223> X=any amino acid

<400> 122

Lys Xaa Ala Thr Gly Ala Phe Leu Ser Ala Glu Arg Gly Gly Lys Met
 1 5 10 15

Ala Val Gln Ile Ser Lys Lys Arg Lys Phe Val Ala Asp Gly Ile Phe
 20 25 30

Lys Ala Glu Leu Asn Glu Phe Leu Thr Arg Glu Leu Ala Glu Asp Gly
 35 40 45

Tyr Ser Gly Val Glu Val Arg Val Thr Pro Thr Arg Thr Glu Ile Ile
 50 55 60

Ile Leu Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg Arg
 65 70 75 80

Ile Arg Glu Leu Thr Ala Val Val Gln Lys Arg Phe Gly Phe Pro Glu
 85 90 95

Gly Ser Val Glu Leu Tyr Ala Glu Lys Val Ala Thr Arg Gly Leu Cys
 100 105 110

Ala Ile Xaa Pro Ala Glu Xaa Leu Xaa Tyr Xaa Leu Xaa Xaa Gly Ser
 115 120 125

151

Leu Arg Arg Val Phe Pro Ile Ala Val Pro His Ala His Gly Ala
130 135 140

<210> 123
<211> 75
<212> PRT
<213> Homo sapien

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (21)..(24)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (26)..(26)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (36)..(36)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (45)..(45)
<223> X=any amino acid

<400> 123

His Xaa Leu Gln Lys His Leu Ala Gly Leu Gly Leu Thr Glu Ala Ile
1 5 10 15

Asp Lys Asn Lys Xaa Xaa Xaa Xaa Arg Xaa Ser Gly Lys Lys Xaa Phe
20 25 30

Tyr Leu Ala Xaa Phe His Ala Thr Ala Phe Glu Leu Xaa Thr Asp Gly
35 40 45

152

Asn Pro Phe Asp Gln Asp Ile Tyr Gly Arg Glu Gly Gly Ala Ala Pro
 50 55 60

Ser Cys Ser Thr Pro Thr Thr Pro Ser Ser Ser
 65 70 75

<210> 124
 <211> 110
 <212> PRT
 <213> Homo sapien

<400> 124

Cys Gly Thr Pro Lys Ala Ala Pro Cys Tyr Ser Leu Gly Ala Trp Ser
 1 5 10 15

Gly Leu Arg Val Thr Arg Cys Glu Thr Ser Tyr Arg Ala Ser Gly Cys
 20 25 30

Thr Gln Asp Gly Arg Arg His Pro Lys Ala Pro Glu Thr His Gly Cys
 35 40 45

Tyr Trp Gly Trp Gly Gly Gly Glu Val Pro Ala Leu Asp Thr Pro Trp
 50 55 60

Gly Gly Gly Gly Lys Thr Asp Arg Gly Ser Arg Val Pro Glu Arg Thr
 65 70 75 80

Phe Pro Ala Arg Ile His Ser Thr Trp Thr Trp Ala Pro Asp Thr Met
 85 90 95

Met Leu Ser Pro Glu Thr Pro His Pro Val Gly Pro Gly Pro
 100 105 110

<210> 125
 <211> 196
 <212> PRT
 <213> Homo sapien

<400> 125

Met Ser Pro Arg Phe Pro Ala Arg Pro Trp Val Val Lys Leu Val Ala
 1 5 10 15

Ser Leu His Glu Asp Leu His Glu Val Ser Val Arg Ser Arg Pro Ser
 20 25 30

Pro Val Pro Thr Pro Gly Trp Leu Gly Glu Gly Val Ala Leu Val Asp
 35 40 45

153

Gly Pro Pro Val Gly Asp Pro Leu Ser Arg Val Pro Glu Pro Cys Arg
 50 55 60

Val Arg Thr Lys Thr Val Lys Lys Ala Ala Arg Val Ile Ile Glu Lys
 65 70 75 80

Tyr Tyr Thr Arg Leu Gly Asn Asp Phe His Thr Asn Lys Arg Val Cys
 85 90 95

Glu Glu Ile Ala Ile Ile Pro Ser Lys Lys Leu Arg Asn Lys Ile Ala
 100 105 110

Gly Tyr Val Thr His Leu Met Lys Arg Ile Gln Arg Gly Pro Val Arg
 115 120 125

Gly Ile Ser Ile Lys Leu Gln Glu Glu Glu Arg Glu Arg Arg Asp Asn
 130 135 140

Tyr Val Pro Glu Val Ser Ala Leu Asp Gln Glu Ile Ile Glu Val Asp
 145 150 155 160

Pro Asp Thr Lys Glu Met Leu Lys Leu Leu Asp Phe Gly Ser Leu Ser
 165 170 175

Asn Leu Gln Val Thr Gln Pro Thr Val Gly Met Asn Phe Lys Thr Pro
 180 185 190

Arg Gly Pro Val
 195

<210> 126
 <211> 207
 <212> PRT
 <213> Homo sapien

<400> 126

Met Pro Glu His Cys Gly Leu Gly His Arg Arg Cys Ala Cys Gln Gln
 1 5 10 15

His Gly Ala Ser Pro Gly Arg Met Thr Phe Glu Gly Asp Thr Asp Val
 20 25 30

Trp Ala Met Pro Gly Ser Trp Glu Gln Arg Pro Arg Ala Gly Pro Gly
 35 40 45

Val Arg Ala Ala Arg Ala Gly Gly Phe Trp Glu Pro Lys Ala Arg Leu

60

Val Arg Thr Lys Thr Val Lys Lys Ala Ala Arg Val Ile Ile Glu Lys
50 55 60

155

Tyr Tyr Thr Arg Leu Gly Asn Asp Phe His Thr Asn Lys Arg Val Cys
 65 70 75 80

Glu Glu Ile Ala Ile Ile Pro Ser Lys Lys Leu Arg Asn Lys Ile Ala
 85 90 95

Gly Tyr Val Thr His Leu Met Lys Arg Ile Gln Arg Gly Pro Val Arg
 100 105 110

Gly Ile Ser Ile Lys Leu Gln Glu Glu Glu Arg Glu Arg Arg Asp Asn
 115 120 125

Tyr Val Pro Glu Val Ser Ala Leu Asp Gln Glu Ile Ile Glu Val Asp
 130 135 140

Pro Asp Thr Lys Glu Met Leu Lys Leu Leu Asp Phe Gly Ser Leu Ser
 145 150 155 160

Asn Leu Gln Val Thr Gln Pro Thr Val Gly Met Asn Phe Lys Thr Pro
 165 170 175

Arg Gly Pro Val
 180

<210> 128
 <211> 150
 <212> PRT
 <213> Homo sapien

<400> 128

Met Gly Arg Val Arg Thr Lys Thr Val Lys Lys Ala Ala Arg Val Ile
 1 5 10 15

Ile Glu Lys Tyr Tyr Thr Arg Leu Gly Asn Asp Phe His Thr Asn Lys
 20 25 30

Arg Val Cys Glu Glu Ile Ala Ile Ile Pro Ser Lys Lys Leu Arg Asn
 35 40 45

Lys Ile Ala Gly Tyr Val Thr His Leu Met Lys Arg Ile Gln Arg Gly
 50 55 60

Pro Val Arg Gly Ile Ser Ile Lys Leu Gln Glu Glu Glu Arg Glu Arg
 65 70 75 80

Arg Asp Asn Tyr Val Pro Glu Val Ser Ala Leu Asp Gln Glu Ile Ile

156

85

90

95

Glu Val Asp Pro Asp Thr Lys Glu Met Leu Lys Leu Leu Asp Phe Gly
 100 105 110

Ser Leu Ser Asn Leu Gln Val Ile His Pro Asn Cys Arg Leu Ser Asp
 115 120 125

Leu Lys Val Gly Gln Thr Ala Gln Pro Thr Val Gly Met Asn Phe Lys
 130 135 140

Thr Pro Arg Gly Pro Val
 145 150

<210> 129
 <211> 298
 <212> PRT
 <213> Homo sapien

<400> 129

Met Arg Leu Ala Ala Leu Ala Val Ser Ala Cys Ile Leu Phe Arg Glu
 1 5 10 15

Ala Leu Leu Arg Pro Trp Thr Gly Pro Pro Glu Arg Met Pro Val Arg
 20 25 30

Ala Ala Arg Gly Glu Gly Pro Val Ala Met Gly Arg Val Ile Arg Gly
 35 40 45

Gln Arg Lys Gly Ala Gly Ser Val Phe Arg Ala His Val Lys His Arg
 50 55 60

Lys Gly Ala Ala Arg Leu Arg Ala Val Asp Phe Ala Glu Arg His Gly
 65 70 75 80

Tyr Ile Lys Gly Ile Val Lys Asp Ile Ile His Asp Pro Gly Arg Gly
 85 90 95

Ala Pro Leu Ala Lys Val Val Phe Arg Asp Pro Tyr Arg Phe Lys Lys
 100 105 110

Arg Thr Glu Leu Phe Ile Ala Ala Glu Gly Ile His Thr Gly Gln Phe
 115 120 125

Val Tyr Cys Gly Lys Lys Ala Gln Leu Asn Ile Gly Asn Val Leu Pro
 130 135 140

157

Val Gly Thr Met Pro Glu Gly Thr Ile Val Cys Cys Leu Glu Glu Lys
 145 150 155 160

Pro Gly Asp Arg Gly Lys Leu Ala Arg Ala Ser Gly Asn Tyr Ala Thr
 165 170 175

Val Ile Ser His Asn Pro Glu Thr Lys Lys Thr Arg Val Lys Leu Pro
 180 185 190

Ser Gly Ser Lys Lys Val Ile Ser Ser Ala Asn Arg Ala Val Val Gly
 195 200 205

Val Val Ala Gly Gly Gly Arg Ile Asp Lys Pro Ile Leu Lys Ala Gly
 210 215 220

Arg Ala Tyr His Lys Tyr Lys Ala Lys Arg Asn Cys Trp Pro Arg Val
 225 230 235 240

Arg Gly Val Ala Met Asn Pro Val Glu His Pro Phe Gly Gly Asn
 245 250 255

His Gln His Ile Gly Lys Pro Ser Thr Ile Arg Arg Asp Ala Pro Ala
 260 265 270

Gly Arg Lys Val Gly Leu Ile Ala Ala Arg Arg Thr Gly Arg Leu Arg
 275 280 285

Gly Thr Lys Thr Val Gln Glu Lys Glu Asn
 290 295

<210> 130
 <211> 271
 <212> PRT
 <213> Homo sapien

<220>
 <221> MISC_FEATURE
 <222> (1)..(2)
 <223> X=any amino acid

<400> 130

Xaa Xaa Ala Gly Ala Gly Ala Arg Gly Glu Gly Pro Val Ala Met Gly
 1 5 10 15

Arg Val Ile Arg Gly Gln Arg Lys Gly Ala Gly Ser Val Phe Arg Ala
 20 25 30

158

His Val Lys His Arg Lys Gly Ala Ala Arg Leu Arg Ala Val Asp Phe
 35 40 45

Ala Glu Arg His Gly Tyr Ile Lys Gly Ile Val Lys Asp Ile Ile His
 50 55 60

Asp Pro Gly Arg Gly Ala Pro Leu Ala Lys Val Val Phe Arg Asp Pro
 65 70 75 80

Tyr Arg Phe Lys Lys Arg Thr Glu Leu Phe Ile Ala Ala Glu Gly Ile
 85 90 95

His Thr Gly Gln Phe Val Tyr Cys Gly Lys Lys Ala Gln Leu Asn Ile
 100 105 110

Gly Asn Val Leu Pro Val Gly Thr Met Pro Glu Gly Thr Ile Val Cys
 115 120 125

Cys Leu Glu Glu Lys Pro Gly Asp Arg Gly Lys Leu Ala Arg Ala Ser
 130 135 140

Gly Asn Tyr Ala Thr Val Ile Ser His Asn Pro Glu Thr Lys Lys Thr
 145 150 155 160

Arg Val Lys Leu Pro Ser Gly Ser Lys Lys Val Ile Ser Ser Ala Asn
 165 170 175

Arg Ala Val Val Gly Val Val Ala Gly Gly Gly Arg Ile Asp Lys Pro
 180 185 190

Ile Leu Lys Ala Gly Arg Ala Tyr His Lys Tyr Lys Ala Lys Arg Asn
 195 200 205

Cys Trp Pro Arg Val Arg Gly Val Ala Met Asn Pro Val Glu His Pro
 210 215 220

Phe Gly Gly Gly Asn His Gln His Ile Gly Lys Pro Ser Thr Ile Arg
 225 230 235 240

Arg Asp Ala Pro Ala Gly Arg Lys Val Gly Leu Ile Ala Ala Arg Arg
 245 250 255

Thr Gly Arg Leu Arg Gly Thr Lys Thr Val Gln Glu Lys Glu Asn
 260 265 270

159

<210> 131
 <211> 550
 <212> PRT
 <213> Homo sapien

<400> 131

Met Met Lys Ala Ala Gly Lys Gln Gln Arg Val Gln Gln Gln His Ser
 1 5 10 15

Ser Ala Gln His Gln Gln His Ala Cys Thr Ala Asn Ser Pro Lys His
 20 25 30

Arg Lys His Val Gly Ser Ser Met Gln Ala Gly Met His Ser Arg Ser
 35 40 45

Gln Ala Ser Ser Thr Ala Gln Gln Gln Leu Lys His Ser Ile Gln Gln
 50 55 60

Gln Gln Ile Pro Leu His Pro Gly Thr Ala Thr Gln Thr Ser Thr Lys
 65 70 75 80

Pro Ile Trp Thr Arg Asn Pro Asp Asp Ile Thr Gln Glu Glu Tyr Gly
 85 90 95

Glu Phe Tyr Lys Ser Leu Thr Asn Asp Trp Glu Asp His Leu Ala Val
 100 105 110

Lys His Phe Ser Val Glu Gly Gln Leu Glu Phe Arg Ala Leu Leu Phe
 115 120 125

Ile Pro Arg Arg Ala Pro Phe Asp Leu Cys Glu Asn Lys Lys Lys Lys
 130 135 140

Asn Asn Ile Lys Leu Tyr Val Arg Arg Val Phe Ile Met Asp Ser Cys
 145 150 155 160

Asp Glu Leu Ile Pro Glu Tyr Leu Asn Phe Ile Arg Gly Val Val Asp
 165 170 175

Ser Glu Asp Leu Pro Leu Asn Ile Ser Arg Glu Met Leu Gln Gln Ser
 180 185 190

Lys Ile Leu Lys Val His Ser Gln Gln Thr Leu Leu Arg Ser Ala Leu
 195 200 205

Ser Ser Ser Leu Glu Leu Ala Glu Asp Lys Ala Glu Leu Gln Asp Asn
 210 215 220

160

Ser Tyr Glu Gly Thr Ser His Lys Asn Leu Asn Ala Trp Asn Pro Arg
 225 230 235 240

Arg His Pro Leu Thr Gly Ala Ala Cys Leu Glu Leu Leu Arg Tyr His
 245 250 255

Thr Ser Gln Ser Gly Asp Glu Met Thr Ser Leu Ser Glu Tyr Val Ser
 260 265 270

Arg Met Lys Glu Thr Gln Lys Ser Ile Tyr Tyr Ile Thr Gly Glu Ser
 275 280 285

Lys Glu Gln Val Ala Asn Ser Ala Phe Val Glu Arg Val Arg Lys Arg
 290 295 300

Gly Phe Glu Val Val Tyr Met Thr Glu Pro Ile Asp Glu Tyr Cys Val
 305 310 315 320

Gln Gln Leu Lys Glu Phe Asp Gly Lys Ser Leu Val Ser Val Thr Lys
 325 330 335

Glu Gly Leu Glu Leu Pro Glu Asp Glu Glu Glu Lys Lys Lys Met Glu
 340 345 350

Glu Ser Lys Ala Lys Phe Glu Asn Leu Cys Lys Leu Met Lys Glu Ile
 355 360 365

Leu Asp Lys Lys Val Glu Lys Val Thr Ile Ser Asn Arg Leu Val Ser
 370 375 380

Ser Pro Cys Cys Ile Val Thr Ser Thr Tyr Gly Trp Thr Ala Asn Met
 385 390 395 400

Glu Arg Ile Met Lys Ala Gln Ala Leu Arg Asp Asn Ser Thr Met Gly
 405 410 415

Tyr Met Met Ala Lys Lys His Leu Glu Ile Asn Pro Asp His Pro Ile
 420 425 430

Val Glu Thr Leu Arg Gln Lys Ala Glu Ala Asp Glu Asn Asp Lys Ala
 435 440 445

Val Lys Asp Leu Val Val Leu Leu Phe Glu Thr Ala Leu Val Ser Ser
 450 455 460

161

Gly Phe Ser Leu Glu Asp Pro Gln Thr Gln Ser Asn Arg Ile Tyr Arg
 465 470 475 480

Met Ile Lys Leu Gly Leu Gly Ile Asp Glu Asp Glu Val Ala Ala Glu
 485 490 495

Glu Pro Asn Ala Ala Val Pro Asp Glu Ile Pro Pro Leu Glu Gly Asp
 500 505 510

Glu Asp Ala Ser Arg Met Arg Gly Ser Arg Val Arg Leu Gly Val Val
 515 520 525

Leu Gly Asn Thr Cys Ala Phe Gly Phe Cys Val Pro His Gly Ala Pro
 530 535 540

Thr Ala Pro Arg Val Pro
 545 550

<210> 132
 <211> 190
 <212> PRT
 <213> Homo sapien

<220>
 <221> MISC_FEATURE
 <222> (181)..(181)
 <223> X=any amino acid

<400> 132

Glu Leu Leu Arg Tyr His Thr Ser Gln Ser Gly Asp Glu Met Thr Ser
 1 5 10 15

Leu Ser Glu Tyr Val Ser Arg Met Lys Glu Thr Gln Lys Ser Ile Tyr
 20 25 30

Tyr Ile Thr Gly Glu Ser Lys Glu Gln Val Ala Asn Ser Ala Phe Val
 35 40 45

Glu Arg Val Arg Lys Arg Gly Phe Glu Val Val Tyr Met Thr Glu Pro
 50 55 60

Ile Asp Glu Tyr Cys Val Gln Gln Leu Lys Glu Phe Asp Gly Lys Ser
 65 70 75 80

Leu Val Ser Val Thr Lys Glu Gly Leu Glu Leu Pro Glu Asp Glu Glu
 85 90 95

162

Glu Lys Lys Lys Met Glu Glu Ser Lys Ala Lys Phe Glu Asn Leu Cys
 100 105 110

Lys Leu Met Lys Glu Ile Leu Asp Lys Lys Val Glu Lys Val Thr Ile
 115 120 125

Ser Asn Arg Leu Val Ser Ser Pro Cys Cys Ile Val Thr Ser Thr Tyr
 130 135 140

Gly Trp Thr Ala Asn Met Glu Arg Ile Met Lys Ala Gln Ala Leu Arg
 145 150 155 160

Asp Asn Ser Thr Met Gly Tyr Met Met Ala Lys Lys His Leu Glu Ile
 165 170 175

Asn Pro Asp His Xaa His Cys Gly Asp Ala Ala Ala Glu Gly
 180 185 190

<210> 133
 <211> 111
 <212> PRT
 <213> Homo sapien

<400> 133

Met Gly Val Asp Ile Arg His Asn Lys Asp Arg Lys Val Arg Arg Lys
 1 5 10 15

Glu Pro Lys Ser Gln Asp Ile Tyr Leu Arg Leu Leu Val Lys Leu Tyr
 20 25 30

Arg Phe Leu Ala Arg Arg Thr Asn Ser Thr Phe Asn Gln Val Val Leu
 35 40 45

Lys Arg Leu Phe Met Ser Arg Thr Asn Arg Pro Pro Leu Ser Leu Ser
 50 55 60

Arg Met Ile Arg Lys Met Lys Leu Pro Gly Arg Glu Asn Lys Thr Ala
 65 70 75 80

Val Val Val Gly Thr Ile Thr Asp Asp Val Arg Val Gln Glu Val Pro
 85 90 95

Arg Arg Asp His Ala Ser Ile Thr Leu Arg Arg Ser Thr Cys Ile
 100 105 110

<210> 134

163

<211> 261

<212> PRT

<213> Homo sapien

<400> 134

Phe Pro Arg Glu Ser Gly Pro Arg Pro Val Pro Arg Thr Asp Ser Gly
1 5 10 15

Ala Ser Val Gly Ala Gly Cys Leu Arg Thr Leu Ala Val Gly Pro Gly
20 25 30

Gln Glu Gly Ala Gly Gly Arg Asp Ser Gly Cys Thr Val Ile Trp Arg
35 40 45

Ser Ala Ala Gly Pro Thr Gly Ile Arg Gly Phe Gly Gly Ala Arg Arg
50 55 60

Pro Gly Ser Glu Leu Gly Ser Cys Cys Ala Ala His Val Leu Thr Ser
65 70 75 80

Ala Ser Asp Val Trp Ser Tyr Gly Ile Val Met Trp Glu Val Met Ser
85 90 95

Tyr Gly Glu Arg Pro Tyr Trp Asp Met Ser Asn Gln Asp Val Ile Asn
100 105 110

Ala Val Glu Gln Asp Tyr Arg Leu Pro Pro Pro Met Asp Cys Pro Thr
115 120 125

Ala Leu His Gln Leu Met Leu Asp Cys Trp Val Arg Asp Arg Asn Leu
130 135 140

Arg Pro Lys Phe Ser Gln Ile Val Asn Thr Leu Asp Lys Leu Ile Arg
145 150 155 160

Asn Ala Ala Ser Leu Lys Val Ile Ala Ser Ala Gln Ser Gly Met Ser
165 170 175

Gln Pro Leu Leu Asp Arg Thr Val Pro Asp Tyr Thr Thr Phe Thr Thr
180 185 190

Val Gly Asp Trp Leu Asp Ala Ile Lys Met Gly Arg Tyr Lys Glu Ser
195 200 205

Phe Val Ser Ala Gly Phe Ala Ser Phe Asp Leu Val Ala Gln Met Thr
210 215 220

164

Ala Glu Asp Leu Leu Arg Ile Gly Val Thr Leu Ala Gly His Gln Lys
 225 230 235 240

Lys Ile Leu Ser Ser Ile Gln Asp Met Arg Leu Gln Met Asn Gln Thr
 245 250 255

Leu Pro Val Gln Val
 260

<210> 135
 <211> 361
 <212> PRT
 <213> Homo sapien

<400> 135

Met Pro Gly Val Cys Asp Arg Ala Pro Asp Phe Leu Ser Pro Ser Glu
 1 5 10 15

Asp Gln Val Leu Arg Pro Ala Leu Gly Ser Ser Val Ala Leu Asn Cys
 20 25 30

Thr Ala Trp Val Val Ser Gly Pro His Cys Ser Leu Pro Ser Val Gln
 35 40 45

Trp Leu Lys Asp Gly Leu Pro Leu Gly Ile Gly Gly His Tyr Ser Leu
 50 55 60

His Glu Tyr Ser Trp Val Lys Ala Asn Leu Ser Glu Val Leu Val Ser
 65 70 75 80

Ser Val Leu Gly Val Asn Val Thr Ser Thr Glu Val Tyr Gly Ala Phe
 85 90 95

Thr Cys Ser Ile Gln Asn Ile Ser Phe Ser Ser Phe Thr Leu Gln Arg
 100 105 110

Ala Gly Pro Thr Ser His Val Ala Ala Val Leu Ala Ser Leu Leu Val
 115 120 125

Leu Leu Ala Leu Leu Leu Ala Ala Leu Leu Tyr Val Lys Cys Arg Leu
 130 135 140

Asn Val Leu Leu Trp Tyr Gln Asp Ala Tyr Gly Glu Val Glu Ile Asn
 145 150 155 160

Asp Gly Lys Leu Tyr Asp Ala Tyr Val Ser Tyr Ser Asp Cys Pro Glu

165
 165 170 175
 Asp Arg Lys Phe Val Asn Phe Ile Leu Lys Pro Gln Leu Glu Arg Arg
 180 185 190
 Arg Gly Tyr Lys Leu Phe Leu Asp Asp Arg Asp Leu Leu Pro Arg Ala
 195 200 205
 Glu Pro Ser Ala Asp Leu Leu Val Asn Leu Ser Arg Cys Arg Arg Leu
 210 215 220
 Ile Val Val Leu Ser Asp Ala Phe Leu Ser Arg Ala Trp Cys Ser His
 225 230 235 240
 Ser Phe Arg Trp Val Pro Arg Gly Val Gly Trp Ala Pro Ala Tyr Thr
 245 250 255
 His Pro Pro Asp Gly Pro Ala Pro Gln Gly Gly Pro Val Pro Ala Ala
 260 265 270
 Gly Ala His Pro Gln Thr His Leu His His Leu Arg Gly Pro Glu Ala
 275 280 285
 Arg Pro Arg Ala Pro Gly Ala Pro Pro Ala Ala Pro Ala Pro Pro Pro
 290 295 300
 Gly Asp Leu Ala Ala Leu Glu Ala Arg Leu Arg Asp Ser Phe Leu Arg
 305 310 315 320
 Phe Leu Glu Arg Ser Ala Ala Gly Ala Ala Ala Glu Gly Ala Val Gln
 325 330 335
 Ala Gly Gly Arg Arg Pro Pro Asp Ala Ala Ala Gly Arg Gln Gly Pro
 340 345 350
 His Ala Asp Ser Ser Arg Pro Ser Pro
 355 360
 <210> 136
 <211> 329
 <212> PRT
 <213> Homo sapien
 <400> 136
 Met Pro Gly Val Cys Asp Arg Ala Pro Asp Phe Leu Ser Pro Ser Glu
 1 5 10 15

166

Asp Gln Val Leu Arg Pro Ala Leu Gly Ser Ser Val Ala Leu Asn Cys
 20 25 30

Thr Ala Trp Val Val Ser Gly Pro His Cys Ser Leu Pro Ser Val Gln
 35 40 45

Trp Leu Lys Asp Gly Leu Pro Leu Gly Ile Gly Gly His Tyr Ser Leu
 50 55 60

His Glu Tyr Ser Trp Val Lys Ala Asn Leu Ser Glu Val Leu Val Ser
 65 70 75 80

Ser Val Leu Gly Val Asn Val Thr Ser Thr Glu Val Tyr Gly Ala Phe
 85 90 95

Thr Cys Ser Ile Gln Asn Ile Ser Phe Ser Ser Phe Thr Leu Gln Arg
 100 105 110

Ala Gly Pro Thr Ser His Val Ala Ala Val Leu Ala Ser Leu Leu Val
 115 120 125

Leu Leu Ala Leu Leu Leu Ala Ala Leu Leu Tyr Val Lys Cys Arg Leu
 130 135 140

Asn Val Leu Leu Trp Tyr Gln Asp Ala Tyr Gly Glu Val Glu Ile Asn
 145 150 155 160

Asp Gly Lys Leu Tyr Asp Ala Tyr Val Ser Tyr Ser Asp Cys Pro Glu
 165 170 175

Asp Arg Lys Phe Val Asn Phe Ile Leu Lys Pro Gln Leu Glu Arg Arg
 180 185 190

Arg Gly Tyr Lys Leu Phe Leu Asp Asp Arg Asp Leu Leu Pro Arg Ala
 195 200 205

Glu Pro Ser Ala Asp Leu Leu Val Asn Leu Ser Arg Cys Arg Arg Leu
 210 215 220

Ile Val Val Leu Ser Asp Ala Phe Leu Ser Arg Ala Trp Cys Ser His
 225 230 235 240

Ser Phe Arg Trp Val Pro Arg Gly Val Gly Trp Ala Pro Ala Tyr Thr
 245 250 255

167

His Pro Pro Asp Gly Pro Ala Pro Gln Gly Gly Pro Val Pro Ala Ala
 260 265 270

Gly Ala His Pro Gln Thr His Leu His His Leu Arg Gly Pro Glu Ala
 275 280 285

Arg Pro Arg Ala Pro Gly Ala Pro Pro Ala Ala Pro Ala Pro Pro Pro
 290 295 300

Gly Asp Leu Ala Ala Leu Glu Ala Arg Leu Arg Gly Ala Glu Gln Ala
 305 310 315 320

Arg Glu Gly Pro Gly Leu Ala Ala Gly
 325

<210> 137
 <211> 164
 <212> PRT
 <213> Homo sapien

<400> 137

Pro Pro Pro Leu Arg Arg Arg Arg Pro Pro Ser Arg Arg Ala Leu Arg
 1 5 10 15

Arg Pro Leu Gly Glu Pro Glu Pro Leu Pro Thr Pro His Arg Gly Ala
 20 25 30

Phe Gly Arg Leu Pro Glu Pro Gly Leu Val Gln Pro Gln Leu Pro Thr
 35 40 45

Pro Ser Ser Asp Phe Trp Lys Glu Val Gln Leu Ala Leu Pro Arg Lys
 50 55 60

Val Arg Tyr Arg Pro Val Glu Gly Asp Pro Gln Thr Gln Leu Gln Asp
 65 70 75 80

Asp Lys Asp Pro Met Leu Ile Leu Arg Gly Arg Val Pro Glu Gly Arg
 85 90 95

Ala Leu Asp Ser Glu Val Asp Pro Asp Pro Glu Gly Asp Leu Gly Val
 100 105 110

Arg Gly Pro Val Phe Gly Glu Pro Ser Ala Pro Pro His Thr Ser Gly
 115 120 125

Val Ser Leu Gly Glu Ser Arg Ser Ser Glu Val Asp Val Ser Asp Leu
 130 135 140

168

Gly Ser Arg Asn Tyr Ser Ala Arg Thr Asp Phe Tyr Cys Leu Val Ser
 145 150 155 160

Lys Asp Asp Met

<210> 138
 <211> 66
 <212> PRT
 <213> Homo sapien

<400> 138

Met Leu Leu Glu Arg Arg Ser Val Met Asp Arg Gly Arg Gly Glu Glu
 1 5 10 15

Trp Arg Ala Arg Ser Glu Ser Ala Gln Ser Lys Met Leu Ser Gly Val
 20 25 30

Gly Gly Phe Val Leu Gly Leu Leu Phe Leu Gly Ala Gly Leu Phe Ile
 35 40 45

Tyr Phe Arg Asn Gln Lys Gly His Ser Gly Leu Gln Pro Thr Gly Phe
 50 55 60

Leu Ser
 65

<210> 139
 <211> 135
 <212> PRT
 <213> Homo sapien

<400> 139

Pro His Ser Arg Lys Asn Leu Leu Pro Gln Leu Cys Arg Met Lys Ser
 1 5 10 15

Phe Pro Ala Trp Gln Leu Phe Phe His Lys Arg Gly Leu Ser Gln Asp
 20 25 30

Leu Val Ala Thr Gly Ser Ala Thr Ala Glu Asn Val Leu Pro Cys Gly
 35 40 45

Phe Leu Ser Ser Cys Pro Trp Pro Glu Val Pro Ala Leu Met Ala Ala
 50 55 60

Pro His Leu Gln Leu Leu Cys Ser Pro Leu Pro Lys Pro Tyr Gly Leu

169

65

70

75

80

Pro Cys Ile Cys Thr His Pro Val Arg Gln Thr His Tyr Ile Ile Lys
85 90 95

Cys Phe Ser Lys Met Glu Leu Asn Ile Ile Trp Ser Ile Trp Leu Gln
100 105 110

Arg Gln Lys Met Lys Arg Lys Arg Glu Asp Tyr Phe Pro Asn Arg Ile
115 120 125

Met Ile Phe Met Tyr Met Ser
130 135

<210> 140
<211> 115
<212> PRT
<213> Homo sapien

<400> 140

Met Lys Ser Phe Pro Ala Trp Gln Leu Phe Phe His Lys Arg Gly Leu
1 5 10 15

Ser Gln Asp Leu Val Ala Thr Gly Ser Ala Thr Leu Gln Lys Met Ser
20 25 30

Ile Pro Cys Gly Phe Leu Ser Ser Cys Pro Trp Pro Glu Val Pro Ala
35 40 45

Leu Met Ala Ala Pro His Leu Gln Leu Leu Cys Ser Pro Leu Pro Lys
50 55 60

Pro Tyr Gly Leu Pro Cys Ile Cys Thr His Pro Val Arg Gln Thr His
65 70 75 80

Tyr Ile Ile Lys Cys Phe Ser Lys Met Glu Leu Asn Ile Ile Trp Ser
85 90 95

Ile Trp Leu Gln Arg Gln Lys Met Lys Arg Lys Arg Glu Asp Tyr Phe
100 105 110

Pro Ile Glu
115

<210> 141
<211> 135
<212> PRT

170

<213> Homo sapien

<400> 141

Pro His Ser Arg Lys Asn Leu Leu Pro Gln Leu Cys Arg Met Lys Ser
 1 5 10 15

Phe Pro Ala Trp Gln Leu Phe Phe His Lys Arg Gly Leu Ser Gln Asp
 20 25 30

Leu Val Ala Thr Gly Ser Ala Thr Ala Glu Asn Val Leu Pro Cys Gly
 35 40 45

Phe Leu Ser Ser Cys Pro Trp Pro Glu Val Pro Ala Leu Met Ala Ala
 50 55 60

Pro His Leu Gln Leu Leu Cys Ser Pro Leu Pro Lys Pro Tyr Gly Leu
 65 70 75 80

Pro Cys Ile Cys Thr His Pro Val Arg Gln Thr His Tyr Ile Ile Lys
 85 90 95

Cys Phe Ser Lys Met Glu Leu Asn Ile Ile Trp Ser Ile Trp Leu Gln
 100 105 110

Arg Gln Lys Met Lys Arg Lys Arg Glu Asp Tyr Phe Pro Asn Arg Ile
 115 120 125

Met Ile Phe Met Tyr Met Ser
 130 135

<210> 142

<211> 220

<212> PRT

<213> Homo sapien

<400> 142

Met Asp Gln His Phe Arg Thr Thr Pro Leu Glu Lys Asn Ala Pro Val
 1 5 10 15

Leu Leu Ala Leu Leu Gly Ile Trp Tyr Ile Asn Cys Phe Gly Cys Glu
 20 25 30

Thr His Ala Met Leu Pro Tyr Asp Gln Tyr Leu His Arg Phe Ala Ala
 35 40 45

Tyr Phe Gln Gln Gly Asp Met Glu Ser Asn Gly Lys Tyr Ile Thr Lys
 50 55 60

171

Ser Gly Thr Arg Val Asp His Gln Thr Gly Pro Ile Val Trp Gly Glu
65 70 75 80

Pro Gly Thr Asn Gly Gln His Ala Phe Tyr Gln Leu Ile His Gln Gly
85 90 95

Thr Lys Met Ile Pro Cys Asp Phe Leu Ile Pro Val Gln Thr Gln His
100 105 110

Pro Ile Arg Lys Gly Leu His His Lys Ile Leu Leu Ala Asn Phe Leu
115 120 125

Ala Gln Thr Glu Ala Leu Met Arg Gly Lys Ser Thr Glu Glu Ala Arg
130 135 140

Lys Glu Leu Gln Ala Ala Gly Lys Ser Pro Glu Asp Leu Glu Arg Leu
145 150 155 160

Leu Pro His Lys Val Phe Glu Gly Asn Arg Pro Thr Asn Ser Ile Val
165 170 175

Phe Thr Lys Leu Thr Pro Phe Met Leu Gly Ala Leu Val Ala Met Tyr
180 185 190

Glu His Lys Ile Phe Val Gln Gly Ile Ile Trp Asp Ile Asn Ser Phe
195 200 205

Asp Gln Trp Gly Ser Gly Ala Gly Lys Ala Ala Gly
210 215 220

<210> 143

<211> 287

<212> PRT

<213> Homo sapien

<220>

<221> MISC_FEATURE

<222> (7)..(7)

<223> X=any amino acid

<400> 143

Val Arg Gly Leu Gly Gly Xaa Ala Ile Gly Leu Ser Ile Ala Leu His
1 5 10 15

Val Gly Phe Asp Asn Phe Glu Gln Leu Leu Ser Gly Ala His Trp Met
20 25 30

172

Asp Gln His Phe Arg Thr Thr Pro Leu Glu Lys Asn Ala Pro Val Leu
 35 40 45

Leu Ala Leu Leu Gly Ile Trp Tyr Ile Asn Cys Phe Gly Cys Glu Thr
 50 55 60

His Ala Met Leu Pro Tyr Asp Gln Tyr Leu His Arg Phe Ala Ala Tyr
 65 70 75 80

Phe Gln Gln Gly Asp Met Glu Ser Asn Gly Lys Tyr Ile Thr Lys Ser
 85 90 95

Gly Thr Arg Val Asp His Gln Thr Gly Pro Ile Val Trp Gly Glu Pro
 100 105 110

Gly Thr Asn Gly Gln His Ala Phe Tyr Gln Leu Ile His Gln Gly Thr
 115 120 125

Lys Met Ile Pro Cys Asp Phe Leu Ile Pro Val Gln Thr Gln His Pro
 130 135 140

Ile Arg Lys Gly Leu His His Lys Ile Leu Leu Ala Asn Phe Leu Ala
 145 150 155 160

Gln Thr Glu Ala Leu Met Arg Gly Lys Ser Thr Glu Glu Ala Arg Lys
 165 170 175

Glu Leu Gln Ala Ala Gly Lys Ser Pro Glu Asp Leu Glu Arg Leu Leu
 180 185 190

Pro His Lys Val Phe Glu Gly Asn Arg Pro Thr Asn Ser Ile Val Phe
 195 200 205

Thr Lys Leu Thr Pro Phe Met Leu Gly Ala Leu Val Ala Met Tyr Glu
 210 215 220

His Lys Ile Phe Val Gln Gly Ile Ile Trp Asp Ile Asn Ser Phe Asp
 225 230 235 240

Gln Trp Gly Val Glu Leu Gly Lys Gln Leu Ala Lys Lys Ile Glu Pro
 245 250 255

Glu Leu Asp Gly Ser Ala Gln Val Thr Ser His Asp Ala Ser Thr Asn
 260 265 270

173

Gly Leu Ile Asn Phe Ile Lys Gln Gln Arg Glu Ala Arg Val Gln
275 280 285

<210> 144
<211> 147
<212> PRT
<213> Homo sapien

<400> 144

Met Ala Pro Gly Arg Gly Leu Gly His Ala Trp Leu Val Leu Gln Asn
1 5 10 15

Gly Arg Ala Cys Pro His Arg Pro Ala Arg Leu Ser Leu Trp Gly Arg
20 25 30

Val Cys Phe Pro Ser Arg Gly Leu Gly Ile Arg Thr Leu Leu Glu Thr
35 40 45

Phe Leu Gly Val Phe Cys Arg Tyr Leu Lys Glu Ile Ala Gln Pro Thr
50 55 60

Leu Leu Cys Ser Pro Ser Ser His His Ser Cys Leu Glu Pro Trp Ser
65 70 75 80

Pro Cys Met Ser Thr Arg Ser Ser Phe Arg Ala Ser Ser Gly Thr Ser
85 90 95

Thr Ala Leu Thr Ser Gly Gly Val Glu Leu Gly Lys Gln Leu Ala Lys
100 105 110

Lys Ile Glu Pro Glu Leu Asp Gly Ser Ala Gln Val Thr Ser His Asp
115 120 125

Ala Ser Thr Asn Gly Leu Ile Asn Phe Ile Lys Gln Gln Arg Glu Ala
130 135 140

Arg Val Gln
145

<210> 145
<211> 150
<212> PRT
<213> Homo sapien

<220>
<221> MISC_FEATURE
<222> (9)..(10)
<223> X=any amino acid

174

<400> 145

Ser Gln His Phe Gly Arg Pro Arg Xaa Xaa Asp His Leu Arg Ser Asp
 1 5 10 15

Gln Ser Gly Gln His Gly Glu Thr Pro Ser Val Pro Lys Ile Gln Lys
 20 25 30

Pro Ala Gly His Gly Gly Thr Cys Leu Trp Ser Gln Leu Leu Gly Arg
 35 40 45

Pro Arg Gln Lys Thr Arg Trp Asn Pro Gly Gly Gly Ala Cys Arg Glu
 50 55 60

Pro Arg Leu Cys His Cys Thr Ala Ala Trp Val Thr Glu Pro Asp Ser
 65 70 75 80

Ile Ser Thr Thr Asp Ala Leu Thr Leu Gly Val Ser Val Ala Gln Gly
 85 90 95

Arg Gly Ala His Val Thr Gln Ala Asp Gly Pro Phe Ala Thr Ala Val
 100 105 110

Asp Glu His Val Ala Leu Val Arg Val Glu Leu Gly Cys Ser Asp Asp
 115 120 125

Phe Gly Gln Leu Leu His Val Ser Arg Leu Asp Val His Asp Val Lys
 130 135 140

Ala Ser Ile Cys Asp Phe
 145 150

<210> 146

<211> 811

<212> PRT

<213> Homo sapien

<400> 146

Met Thr Asp Ile Leu Phe Leu Pro Met Trp Ile Ser Asn Gln His Thr
 1 5 10 15

Pro Ser Ser Pro Gln Gly Asp Gly Gly Ser Ala His Thr Phe Ile Ser
 20 25 30

Thr Gly Gly Pro Gly Ile Ser Thr Arg Leu His Leu His Arg Gly Met
 35 40 45

175

Gly Asp Gln His Thr Pro Ser Ser Pro Gln Trp Asp Gly Gly Ser Ala
 50 55 60

His Ala Phe Ile Ser Thr Gly Gly Trp Gly Met Ser Thr Arg Leu His
 65 70 75 80

Leu His Arg Gly Met Ala Asp Gln His Thr Pro Ser Ser Pro Gln Gly
 85 90 95

Asp Gly Gly Ser Ala His Ala Phe Ile Ser Thr Gly Gly Arg Gly Ile
 100 105 110

Ser Thr Arg Leu His Leu His Arg Arg Thr Gly Asp Gln His Thr Pro
 115 120 125

Ser Ser Pro Gln Gly Asp Arg Gly Ser Ala His Thr Phe Ile Ser Thr
 130 135 140

Gly Gly Trp Gly Ile Ser Thr His Leu His Leu His Arg Gly Met Gly
 145 150 155 160

Asp Gln His Thr Pro Ser Ser Pro Gln Gly Asp Gly Gly Ser Ala His
 165 170 175

Ala Phe Ile Ser Thr Gly Gly Trp Gly Ile Ser Thr Arg Leu His Leu
 180 185 190

His Ser Gly Met Ala Asp Gln His Thr Pro Ser Ser Pro Gln Gly Asp
 195 200 205

Gly Gly Ser Ala His Thr Phe Ile Ser Thr Gly Gly Trp Gly Ile Ser
 210 215 220

Thr Arg Leu His Leu His Ser Gly Met Ala Asp Gln His Thr Pro Ser
 225 230 235 240

Ser Pro Gln Gly Asp Gly Gly Ser Ala His Ala Phe Ile Ser Thr Val
 245 250 255

Gly Arg Gly Ile Ser Thr His Leu His Leu His Arg Gly Thr Gly Asp
 260 265 270

Gln His Thr Pro Pro Ser Pro Gln Gly His Glu Glu Ala Ala His Thr
 275 280 285

176

Phe Ile Ser Thr Gly Gly Arg Gly Ile Ser Thr His Leu His Leu His
 290 295 300

Arg Gly Met Gly Asp Gln His Thr Pro Pro Ser Pro Gln Gly Asp Lys
 305 310 315 320

Arg Ser Ala His Thr Phe Ile Pro Thr Gly Gly Gln Gly Ile Ser Ile
 325 330 335

Pro Leu His Leu His Arg Gly Met Gly Asp Gln His Thr Pro Ser Ser
 340 345 350

Pro Gln Gly Asp Gly Gly Ser Ala Tyr Pro Phe Ile Ser Thr Gly Gly
 355 360 365

Trp Gly Ile Ser Thr His Leu His Pro His Arg Gly Met Gly Asp Gln
 370 375 380

His Thr Pro Pro Ser Pro Gln Gly His Glu Glu Ser Ala His Thr Phe
 385 390 395 400

Ile Ser Thr Gly Arg Arg Gly Ile Ser Thr Pro Leu His Leu His Arg
 405 410 415

Gly Met Gly Asp Gln His Thr Pro Ser Ser Pro Gln Gly Asp Gly Gly
 420 425 430

Ser Ala Val His Thr Phe Ile Lys Ile Gly Glu Gln Gly Ile Ser Thr
 435 440 445

His Leu Tyr Leu His Arg Gly Thr Arg Asp Gln His Thr Pro Pro Ser
 450 455 460

Pro Gln Gly Met Gly Asp Gln His Thr Pro Ser Ser Pro Gln Gly Asp
 465 470 475 480

Gly Asp Gln His Thr Pro Ser Ser Pro Gln Gly Asp Gly Gly Ser Thr
 485 490 495

His Pro Phe Ile Ser Thr Gly Asp Gly Gly Ser Ala His Thr Phe Ile
 500 505 510

Ser Thr Gly Gly Arg Gly Ile Ser Thr Arg Leu His Val His Arg Gly
 515 520 525

177

Thr Gly Asp Gln His Thr Pro Ser Ser Ser Gln Gly Asp Gly Gly Ser
 530 535 540

Ala His Thr Phe Ile Ser Thr Gly Gly Arg Gly Ser Ala His Thr Ile
 545 550 555 560

Ser Thr Gly Gly Gln Gly Ile Asn Thr Pro Leu His Leu His Met Gly
 565 570 575

Met Gly Asp Gln His Thr Pro Ser Ser Pro Gln Gly Asp Gly Asp Gln
 580 585 590

His Thr Pro Pro Ser Pro Gln Gly Arg Gly Gly Leu Ala His Pro Phe
 595 600 605

Ile Ser Thr Gly Arg Trp Gly Ile Ser Thr His Leu His Leu His Arg
 610 615 620

Gly Thr Gly Asp Gln His Thr Pro Ser Ser Pro Gln Trp Asp Arg Gly
 625 630 635 640

Ser Ala Tyr Pro Phe Ile Ser Thr Gly Gly Trp Gly Ser Ala His Thr
 645 650 655

Phe Ile Ser Thr Glu Glu Met Gly Asp Gln His Ala Pro Ser Ser Pro
 660 665 670

Gln Gly His Gly Gly Ser Ala His Thr Phe Ile Ser Thr Gly Gly Arg
 675 680 685

Gly Ile Ser Thr His Leu His Pro Asp Arg Gly Met Arg Asn Gln His
 690 695 700

Thr Pro Ser Ser Arg Gln Gly Asp Gly Met Gly Asp Gln His Thr Pro
 705 710 715 720

Pro Ser Pro Gln Gly His Glu Gly Ala Ala His Thr Ser Ile Ser Thr
 725 730 735

Gly His Arg Gly Ser Ala His Thr Ser Phe Ser Thr Gly Ala Gln Ala
 740 745 750

Ile Ser Thr Tyr Leu His Leu Asp Arg Val Thr Gly Asp Gln His Thr
 755 760 765

Pro Pro Ser Pro Gln Gln Gln Glu Glu Ser Thr His Thr Phe Ile Ser

178

770

775

780

Thr Gly Gly Arg Gly Ile Ser Thr His Leu His Leu His Arg Gly Thr
 785 790 795 800

Gly Ala Arg Leu Pro Thr Pro Leu Gly Asp Thr
 805 810

<210> 147
 <211> 442
 <212> PRT
 <213> Homo sapien

<400> 147

Phe Arg Val Met Thr Asp Ile Leu Phe Leu Pro Met Trp Ile Ser Asn
 1 5 10 15

Gln His Thr Pro Ser Ser Pro Gln Gly Asp Gly Gly Ser Ala His Thr
 20 25 30

Phe Ile Ser Thr Gly Gly Pro Gly Ile Ser Thr Arg Leu His Leu His
 35 40 45

Arg Gly Met Gly Asp Gln His Thr Pro Ser Ser Pro Gln Trp Asp Gly
 50 55 60

Gly Ser Ala His Ala Phe Ile Ser Thr Gly Gly Trp Gly Met Ser Thr
 65 70 75 80

Arg Leu His Leu His Arg Gly Met Ala Asp Gln His Thr Pro Ser Ser
 85 90 95

Pro Gln Gly Asp Gly Gly Ser Ala His Ala Phe Ile Ser Thr Gly Gly
 100 105 110

Arg Gly Ile Ser Thr Arg Leu His Leu His Arg Arg Thr Gly Asp Gln
 115 120 125

His Thr Pro Ser Ser Pro Gln Gly Asp Arg Gly Ser Ala His Thr Phe
 130 135 140

Ile Ser Thr Gly Gly Trp Gly Ile Ser Thr His Leu His Leu His Arg
 145 150 155 160

Gly Met Gly Asp Gln His Thr Pro Ser Ser Pro Gln Gly Asp Gly Gly
 165 170 175

179

Ser Ala His Ala Phe Ile Ser Thr Gly Gly Trp Gly Ile Ser Thr Arg
 180 185 190

Leu His Leu His Ser Gly Met Ala Asp Gln His Thr Pro Ser Ser Pro
 195 200 205

Gln Gly Asp Gly Gly Ser Ala His Thr Phe Ile Ser Thr Gly Gly Trp
 210 215 220

Gly Ile Ser Thr Arg Leu His Leu His Ser Gly Met Ala Asp Gln His
 225 230 235 240

Thr Pro Ser Ser Pro Gln Gly Asp Gly Gly Ser Ala His Ala Phe Ile
 245 250 255

Ser Thr Val Gly Arg Gly Ile Ser Thr His Leu His Leu His Arg Gly
 260 265 270

Thr Gly Asp Gln His Thr Pro Pro Ser Pro Gln Gly His Glu Glu Ala
 275 280 285

Ala His Thr Phe Ile Ser Thr Gly Gly Arg Gly Ile Ser Thr His Leu
 290 295 300

His Leu His Arg Gly Met Gly Asp Gln His Thr Pro Pro Ser Pro Gln
 305 310 315 320

Gly Asp Lys Arg Ser Ala His Thr Phe Ile Pro Thr Gly Gly Gln Gly
 325 330 335

Ile Ser Ile Pro Leu His Leu His Arg Gly Met Gly Asp Gln His Thr
 340 345 350

Pro Ser Ser Pro Gln Gly Asp Gly Gly Ser Ala Tyr Pro Phe Ile Ser
 355 360 365

Thr Gly Gly Trp Gly Ile Ser Thr His Leu His Pro His Arg Gly Met
 370 375 380

Gly Asp Gln His Thr Pro Pro Ser Pro Gln Gly His Glu Glu Ser Ala
 385 390 395 400

His Thr Phe Ile Ser Thr Gly Arg Arg Gly Ile Ser Thr Pro Leu His
 405 410 415

180

Leu His Arg Gly Met Gly Asp Gln His Thr Pro Ser Ser Pro Gln Gly
 420 425 430

Asp Gly Gly Ser Ala Val His Thr Phe Ile
 435 440

<210> 148

<211> 351

<212> PRT

<213> Homo sapien

<400> 148

Met Lys Ala Ser Gly Thr Leu Arg Glu Tyr Lys Val Val Gly Arg Cys
 1 5 10 15

Leu Pro Thr Pro Lys Cys His Thr Pro Pro Leu Tyr Arg Met Arg Ile
 20 25 30

Phe Ala Pro Asn His Val Val Ala Lys Ser Arg Phe Trp Tyr Phe Val
 35 40 45

Ser Gln Leu Lys Lys Met Lys Lys Ser Ser Gly Glu Ile Val Tyr Cys
 50 55 60

Gly Gln Val Phe Glu Lys Ser Pro Leu Arg Val Lys Asn Phe Gly Ile
 65 70 75 80

Trp Leu Arg Tyr Asp Ser Arg Ser Gly Thr His Asn Met Tyr Arg Glu
 85 90 95

Tyr Arg Asp Leu Thr Thr Ala Gly Ala Val Thr Gln Cys Tyr Arg Asp
 100 105 110

Met Gly Ala Arg His Arg Ala Arg Ala His Ser Ile Gln Ile Met Lys
 115 120 125

Val Glu Glu Ile Ala Ala Ser Lys Cys Arg Arg Pro Ala Val Lys Gln
 130 135 140

Phe His Asp Ser Lys Ile Lys Phe Pro Leu Pro His Arg Val Leu Arg
 145 150 155 160

Arg Gln His Lys Pro Arg Phe Thr Thr Lys Arg Pro Asn Asn Leu Leu
 165 170 175

Ser Arg Cys Arg Ala Leu Val Arg Gly Val Pro Pro Asn Lys Leu Arg
 180 185 190

181

Asn Ala Pro Lys Val Gln Ala Ala Val Pro Asp Asp Phe Lys Asp Phe
 195 200 205

Ser Leu Leu Asn Glu Glu Ala Arg Tyr Tyr Gln Phe Lys Thr Met Val
 210 215 220

Arg Arg Ala Trp Ser Ala Gly Thr His Asp Pro Glu Lys Ser Thr Gly
 225 230 235 240

Asn Arg Asp Gly Glu Arg Leu Asp Ala Lys Ser Ser Ala Arg Arg Trp
 245 250 255

Ala Lys Arg Asp Arg Thr Thr Arg Arg Ala Leu Pro Ala Glu Glu Glu
 260 265 270

Tyr His Ser Asn Ala Lys Ala Thr Val Arg Gln Asn Lys Pro Arg Arg
 275 280 285

His Gln Ser Gly Ala Lys Glu Lys Lys Gln His Asn Glu His Ala Ala
 290 295 300

Ala Gln Tyr Ala Ala Arg Ser Lys Glu Thr Asp Arg Lys Gln Pro Val
 305 310 315 320

Gly Asp Asn Gln Gly Glu Thr Lys Pro Pro Gly Arg Lys Arg Glu Gly
 325 330 335

Glu Glu Arg Thr Ala Gly Pro Asn Lys Glu Arg Asn Ser Arg His
 340 345 350

<210> 149

<211> 223

<212> PRT

<213> Homo sapien

<220>

<221> MISC_FEATURE

<222> (4)..(4)

<223> X=any amino acid

<400> 149

Ala Phe Ala Xaa Gly Gly Glu Arg Gly Glu His Ala Met Lys Ala Ser
 1 5 10 15

Gly Thr Leu Arg Glu Tyr Lys Val Val Gly Arg Cys Leu Pro Thr Pro
 20 25 30

Lys Cys His Thr Pro Pro Leu Tyr Arg Met Arg Ile Phe Ala Pro Asn
35 40 45

His Val Val Ala Lys Ser Arg Phe Trp Tyr Phe Val Ser Gln Leu Lys
50 55 60

Lys Met Lys Lys Ser Ser Gly Glu Ile Val Tyr Cys Gly Gln Val Phe
65 70 75 80

Glu Lys Ser Pro Leu Arg Val Lys Asn Phe Gly Ile Trp Leu Arg Tyr
85 90 95

Asp Ser Arg Ser Gly Thr His Asn Met Tyr Arg Glu Tyr Arg Asp Leu
100 105 110

Thr Thr Ala Gly Ala Val Thr Gln Cys Tyr Arg Asp Met Gly Ala Arg
115 120 125

His Arg Ala Arg Ala His Ser Ile Gln Ile Met Lys Val Glu Glu Ile
130 135 140

Ala Ala Ser Lys Cys Arg Arg Pro Ala Val Lys Gln Phe His Asp Ser
145 150 155 160

Lys Ile Lys Phe Pro Leu Pro His Arg Val Leu Arg Arg Gln His Lys
165 170 175

Pro Arg Phe Thr Thr Lys Arg Pro Asn Asn Leu Leu Ser Arg Cys Arg
180 185 190

Ala Leu Val Arg Gly Val Pro Pro Asn Lys Leu Arg Asn Ala Pro Lys
195 200 205

Val Gln Ala Ala Val Pro Asp Asp Phe Lys Asp Phe Ser Leu Leu
210 215 220

<210>	150
<211>	260
<212>	PRT
<213>	Homo sapien

<400> 150

Thr Ala Val Leu Ser Pro Gly Pro Arg Leu Pro Ser His Ser Ala Arg
1 5 10 15

183

Cys Ala Cys Glu Gly Leu Ala Ala Leu Gly Thr Gly Gly Ala Ala Arg
 20 25 30

Gly Val Arg Val Gly Val Arg Glu Gly Ser Thr Gln Asp Leu Arg Thr
 35 40 45

Leu Leu Trp Gly Arg Thr Lys His Leu Pro Gly Ala Gly Gly Ala Pro
 50 55 60

Gly Thr Arg Arg Phe Arg Gln Leu Gly Ala Leu Gly Ile Cys Gly Leu
 65 70 75 80

Arg Pro Gly Asp Gly Leu Gly Gly His Ala His Ala Leu Gly Leu Thr
 85 90 95

Glu Cys Asp Arg Ala Arg Gly Arg Ala Lys Arg Gly Gly Arg Ala Arg
 100 105 110

Arg Arg Lys Glu Gly Leu Val Arg Pro Ala Gln Pro Asp Gln Cys Arg
 115 120 125

Gly Gly Asn Gly Leu Gly Ala Gly Pro Ile Arg Ala Gly Gly Phe Leu
 130 135 140

Arg Arg Arg Pro Ser Pro Gln Leu Leu Asp Cys Ser Gly Ala Gly Gly
 145 150 155 160

Thr Asn Thr Trp Arg Phe Phe Arg Arg Gly Glu Asp Phe Leu Arg Ala
 165 170 175

Gln Arg Ile His Phe Leu His Ile Asn Leu Ser Cys Trp Arg Asp Thr
 180 185 190

Ala Gly Lys Arg Arg Pro Ile Phe Val Gln Arg Thr Leu Asp Leu Gly
 195 200 205

Arg Asn Lys Asp Asp Leu Asp Pro Cys Pro His Tyr Leu Glu Phe Ser
 210 215 220

Met Leu Ala Lys Ile Trp Thr Arg Ala Val Pro Glu Gly Arg Gly Pro
 225 230 235 240

Trp Arg Glu Ala Pro Val Thr Ala His Pro Gly Val Gly Leu Trp Ala
 245 250 255

Leu Leu Leu Cys

184

260

<210> 151
 <211> 259
 <212> PRT
 <213> Homo sapien

<400> 151

Ser Arg Val Val Ala Arg Pro Arg Leu Pro Ser His Ser Ala Arg Cys
 1 5 10 15

Ala Cys Glu Gly Leu Ala Ala Leu Gly Thr Gly Gly Ala Ala Arg Gly
 20 25 30

Val Arg Val Gly Val Arg Glu Gly Ser Thr Gln Asp Leu Arg Thr Leu
 35 40 45

Leu Trp Gly Arg Thr Lys His Leu Pro Gly Ala Gly Gly Ala Pro Gly
 50 55 60

Thr Arg Arg Phe Arg Gln Leu Gly Ala Leu Gly Ile Cys Gly Leu Arg
 65 70 75 80

Pro Gly Asp Gly Leu Gly Gly His Ala His Ala Leu Gly Leu Thr Glu
 85 90 95

Cys Asp Arg Ala Arg Gly Arg Ala Lys Arg Gly Gly Arg Ala Arg Arg
 100 105 110

Arg Lys Glu Gly Leu Val Arg Pro Ala Gln Pro Asp Gln Cys Arg Gly
 115 120 125

Gly Asn Gly Leu Gly Ala Gly Pro Ile Arg Ala Gly Gly Phe Leu Arg
 130 135 140

Arg Arg Pro Ser Pro Gln Leu Leu Asp Cys Ser Gly Ala Gly Gly Thr
 145 150 155 160

Asn Thr Trp Arg Phe Phe Arg Arg Gly Glu Asp Phe Leu Arg Ala Gln
 165 170 175

Arg Ile His Phe Leu His Ile Asn Leu Ser Cys Trp Arg Asp Thr Ala
 180 185 190

Gly Lys Arg Arg Pro Ile Phe Val Gln Arg Thr Leu Asp Leu Gly Arg
 195 200 205

185

Asn Lys Asp Asp Leu Asp Pro Cys Pro His Tyr Leu Glu Phe Ser Met
 210 215 220

Leu Ala Lys Ile Trp Thr Arg Ala Val Pro Glu Gly Arg Gly Pro Trp
 225 230 235 240

Arg Glu Ala Pro Val Thr Ala His Pro Gly Val Gly Leu Trp Ala Leu
 245 250 255

Leu Leu Cys

<210> 152
 <211> 650
 <212> PRT
 <213> Homo sapien

<400> 152

Met Gln Gln Asp Gly Leu Gly Val Gly Thr Arg Asn Gly Ser Gly Lys
 1 5 10 15

Gly Arg Ser Val His Pro Ser Trp Pro Trp Cys Ala Pro Arg Pro Leu
 20 25 30

Arg Tyr Phe Gly Arg Asp Ala Arg Ala Arg Arg Ala Gln Thr Ala Ala
 35 40 45

Met Ala Leu Leu Ala Gly Gly Leu Ser Arg Gly Leu Gly Ser His Pro
 50 55 60

Ala Ala Ala Gly Arg Asp Ala Val Val Phe Val Trp Leu Leu Leu Ser
 65 70 75 80

Thr Trp Cys Thr Ala Pro Ala Arg Ala Ile Gln Val Thr Val Ser Asn
 85 90 95

Pro Tyr His Val Val Ile Leu Phe Gln Pro Val Thr Leu Pro Cys Thr
 100 105 110

Tyr Gln Met Thr Ser Thr Pro Thr Gln Pro Ile Val Ile Trp Lys Tyr
 115 120 125

Lys Ser Phe Cys Arg Asp Arg Ile Ala Asp Ala Phe Ser Pro Ala Ser
 130 135 140

Val Asp Asn Gln Leu Asn Ala Gln Leu Ala Ala Gly Asn Pro Gly Tyr

186

145		150		155		160
Asn Pro Tyr Val Glu Cys Gln Asp Ser Val Arg Thr Val Arg Val Val	165		170		175	
Ala Thr Lys Gln Gly Asn Ala Val Thr Leu Gly Asp Tyr Tyr Gln Gly	180		185		190	
Arg Arg Ile Thr Ile Thr Gly Asn Ala Asp Leu Thr Phe Asp Gln Thr	195		200		205	
Ala Trp Gly Asp Ser Gly Val Tyr Tyr Cys Ser Val Val Ser Ala Gln	210		215		220	
Asp Leu Gln Gly Asn Asn Glu Ala Tyr Ala Glu Leu Ile Val Leu Gly	225		230		235	240
Arg Thr Ser Gly Val Ala Glu Leu Leu Pro Gly Phe Gln Ala Gly Pro	245		250		255	
Ile Glu Asp Trp Leu Phe Val Val Val Val Cys Leu Ala Ala Phe Leu	260		265		270	
Ile Phe Leu Leu Leu Gly Ile Cys Trp Cys Gln Cys Cys Pro His Thr	275		280		285	
Cys Cys Cys Tyr Val Arg Cys Pro Cys Cys Pro Asp Lys Cys Cys Cys	290		295		300	
Pro Glu Ala Leu Tyr Ala Ala Gly Lys Ala Ala Thr Ser Gly Val Pro	305		310		315	320
Ser Ile Tyr Ala Pro Ser Thr Tyr Ala His Leu Ser Pro Ala Lys Thr	325		330		335	
Pro Pro Pro Pro Ala Met Ile Pro Met Gly Pro Ala Tyr Asn Gly Tyr	340		345		350	
Pro Gly Gly Tyr Pro Gly Asp Val Asp Arg Ser Ser Ser Ala Gly Gly	355		360		365	
Gln Gly Ser Tyr Val Pro Leu Leu Arg Asp Thr Asp Ser Ser Val Ala	370		375		380	
Ser Glu Val Arg Ser Gly Tyr Arg Ile Gln Ala Ser Gln Gln Asp Asp	385		390		395	400

187

Ser Met Arg Val Leu Tyr Tyr Met Glu Lys Glu Leu Ala Asn Phe Asp
 405 410 415

Pro Ser Arg Pro Gly Pro Pro Ser Gly Arg Val Glu Arg Ala Met Ser
 420 425 430

Glu Val Thr Ser Leu His Glu Asp Asp Trp Arg Ser Arg Pro Ser Arg
 435 440 445

Gly Pro Ala Leu Thr Pro Ile Arg Asp Glu Glu Trp Gly Gly His Ser
 450 455 460

Pro Arg Ser Pro Arg Gly Trp Asp Gln Glu Pro Ala Arg Glu Gln Ala
 465 470 475 480

Gly Gly Gly Trp Arg Ala Arg Arg Pro Arg Ala Arg Ser Val Asp Ala
 485 490 495

Leu Asp Asp Leu Thr Pro Pro Ser Thr Ala Glu Ser Gly Ser Arg Ser
 500 505 510

Pro Thr Ser Asn Gly Gly Arg Arg Ser Arg Ala Tyr Met Pro Pro Arg
 515 520 525

Ser Arg Ser Arg Asp Asp Leu Tyr Asp Gln Asp Asp Ser Arg Asp Phe
 530 535 540

Pro Arg Ser Arg Asp Pro His Tyr Asp Asp Phe Arg Ser Arg Glu Arg
 545 550 555 560

Pro Pro Ala Asp Pro Arg Ser His His His Arg Thr Arg Asp Pro Arg
 565 570 575

Asp Asn Gly Ser Arg Ser Gly Asp Leu Pro Tyr Asp Gly Arg Leu Leu
 580 585 590

Glu Glu Ala Val Arg Lys Lys Gly Ser Glu Glu Arg Arg Arg Pro His
 595 600 605

Lys Glu Glu Glu Glu Glu Ala Tyr Tyr Pro Pro Ala Pro Pro Pro Tyr
 610 615 620

Ser Glu Thr Asp Ser Gln Ala Ser Arg Glu Arg Arg Leu Lys Lys Asn
 625 630 635 640

188

Leu Ala Leu Ser Arg Glu Ser Leu Val Val
 645 650

<210> 153
 <211> 388
 <212> PRT
 <213> Homo sapien

<400> 153

Met Ser Lys Glu Ala Leu Gln Arg Arg Gly Arg Leu Gly Lys Glu Val
 1 5 10 15

Gln Ala Gln Val Pro Pro Glu Pro Asn Gly Tyr Gly Ala Ala Trp Leu
 20 25 30

Leu Pro His Pro Pro Ser Pro Val Asp Cys Val Leu Thr Val Tyr Ala
 35 40 45

Ala Gly Lys Ala Ala Thr Ser Gly Val Pro Ser Ile Tyr Ala Pro Ser
 50 55 60

Thr Tyr Ala His Leu Ser Pro Ala Lys Thr Pro Pro Pro Pro Ala Met
 65 70 75 80

Ile Pro Met Gly Pro Ala Tyr Asn Gly Tyr Pro Gly Gly Tyr Pro Gly
 85 90 95

Asp Val Asp Arg Ser Ser Ser Ala Gly Gly Gln Gly Ser Tyr Val Pro
 100 105 110

Leu Leu Arg Asp Thr Asp Ser Ser Val Ala Ser Glu Val Arg Ser Gly
 115 120 125

Tyr Arg Ile Gln Ala Ser Gln Gln Asp Asp Ser Met Arg Val Leu Tyr
 130 135 140

Tyr Met Glu Lys Glu Leu Ala Asn Phe Asp Pro Ser Arg Pro Gly Pro
 145 150 155 160

Pro Ser Gly Arg Val Glu Arg Ala Met Ser Glu Val Thr Ser Leu His
 165 170 175

Glu Asp Asp Trp Arg Ser Arg Pro Ser Arg Gly Pro Ala Leu Thr Pro
 180 185 190

Ile Arg Asp Glu Glu Trp Gly Gly His Ser Pro Arg Ser Pro Arg Gly

189

195

200

205

Trp Asp Gln Glu Pro Ala Arg Glu Gln Ala Gly Gly Gly Trp Arg Ala
 210 215 220

Arg Arg Pro Arg Ala Arg Ser Val Asp Ala Leu Asp Asp Leu Thr Pro
 225 230 235 240

Pro Ser Thr Ala Glu Ser Gly Ser Arg Ser Pro Thr Ser Asn Gly Gly
 245 250 255

Arg Arg Ser Arg Ala Tyr Met Pro Pro Arg Ser Arg Ser Arg Asp Asp
 260 265 270

Leu Tyr Asp Gln Asp Asp Ser Arg Asp Phe Pro Arg Ser Arg Asp Pro
 275 280 285

His Tyr Asp Asp Phe Arg Ser Arg Glu Arg Pro Pro Ala Asp Pro Arg
 290 295 300

Ser His His His Arg Thr Arg Asp Pro Arg Asp Asn Gly Ser Arg Ser
 305 310 315 320

Gly Asp Leu Pro Tyr Asp Gly Arg Leu Leu Glu Glu Ala Val Arg Lys
 325 330 335

Lys Gly Ser Glu Glu Arg Arg Arg Pro His Lys Glu Glu Glu Glu
 340 345 350

Ala Tyr Tyr Pro Pro Ala Pro Pro Pro Tyr Ser Glu Thr Asp Ser Gln
 355 360 365

Ala Ser Arg Glu Arg Arg Leu Lys Lys Asn Leu Ala Leu Ser Arg Glu
 370 375 380

Ser Leu Val Val
 385

<210> 154
 <211> 83
 <212> PRT
 <213> Homo sapien

<400> 154

Met Lys Pro Gly Glu Gly Gly Gln Val Ala Pro Ser Leu Pro Gly Ser
 1 5 10 15

190

Gly Gln Thr Cys Leu Glu Ser Gln Gly Arg Thr Arg Ser Ser Asn Pro
 20 25 30

Pro Thr Ala Pro Ser Arg Leu Pro Ala Arg Pro Thr Ser His Ser Leu
 35 40 45

Gly Ser His Gly Ala Asp Arg Pro Arg Arg Glu His Thr Pro Pro Val
 50 55 60

Cys Ala Leu Ser Arg Ser Gln Arg Pro Arg Gly His Arg Ala Met His
 65 70 75 80

Ala Pro Asn

<210> 155
 <211> 379
 <212> PRT
 <213> Homo sapien

<400> 155

Ala Ser His Leu Leu Pro Gln Ala Pro Thr Ala Ser Pro Cys Val Leu
 1 5 10 15

Gln Glu Thr Tyr Lys Leu Pro His Arg Leu Ile Glu Lys Lys Arg Arg
 20 25 30

Asp Arg Ile Asn Glu Cys Ile Ala Gln Leu Lys Asp Leu Leu Pro Glu
 35 40 45

His Leu Lys Leu Thr Thr Leu Gly His Leu Glu Lys Ala Val Val Leu
 50 55 60

Glu Leu Thr Leu Lys His Val Lys Ala Leu Thr Asn Leu Ile Asp Gln
 65 70 75 80

Gln Gln Gln Lys Ile Ile Ala Leu Gln Ser Gly Leu Gln Ala Gly Glu
 85 90 95

Leu Ser Gly Arg Asn Val Glu Thr Gly Gln Glu Met Phe Cys Ser Gly
 100 105 110

Phe Gln Thr Cys Ala Arg Glu Val Leu Gln Tyr Leu Ala Lys His Glu
 115 120 125

Asn Thr Arg Asp Leu Lys Ser Ser Gln Leu Val Thr His Leu His Arg

191

130

135

140

Val Val Ser Glu Leu Leu Gln Gly Gly Thr Ser Arg Lys Pro Ser Asp
 145 150 155 160

Pro Ala Pro Lys Val Met Asp Phe Lys Glu Lys Pro Ser Ser Pro Ala
 165 170 175

Lys Gly Ser Glu Gly Pro Gly Lys Asn Cys Val Pro Val Ile Gln Arg
 180 185 190

Thr Phe Ala His Ser Ser Gly Glu Gln Ser Gly Ser Asp Thr Asp Thr
 195 200 205

Asp Ser Gly Tyr Gly Gly Glu Ser Glu Lys Gly Asp Leu Arg Ser Glu
 210 215 220

Gln Pro Cys Phe Lys Ser Asp His Gly Arg Arg Phe Thr Met Gly Glu
 225 230 235 240

Arg Ile Gly Ala Ile Lys Gln Glu Ser Glu Glu Pro Pro Thr Lys Lys
 245 250 255

Asn Arg Met Gln Leu Ser Asp Asp Glu Gly His Phe Thr Ser Ser Asp
 260 265 270

Leu Ile Ser Ser Pro Phe Leu Gly Pro His Pro His Gln Pro Pro Phe
 275 280 285

Cys Leu Pro Phe Tyr Leu Ile Pro Pro Ser Ala Thr Ala Tyr Leu Pro
 290 295 300

Met Leu Glu Lys Cys Trp Tyr Pro Thr Ser Val Pro Val Leu Tyr Pro
 305 310 315 320

Gly Leu Asn Ala Ser Ala Ala Ala Leu Ser Ser Phe Met Asn Pro Asp
 325 330 335

Lys Ile Ser Ala Pro Leu Leu Met Pro Gln Arg Leu Pro Ser Pro Leu
 340 345 350

Pro Ala His Pro Ser Val Asp Ser Ser Val Leu Leu Gln Ala Leu Lys
 355 360 365

Pro Ile Pro Pro Leu Asn Leu Glu Thr Lys Asp
 370 375

192

<210> 156
 <211> 379
 <212> PRT
 <213> Homo sapien

<400> 156

Ala Ser His Leu Leu Pro Gln Ala Pro Thr Ala Ser Pro Cys Val Leu
 1 5 10 15

Gln Glu Thr Tyr Lys Leu Pro His Arg Leu Ile Glu Lys Lys Arg Arg
 20 25 30

Asp Arg Ile Asn Glu Cys Ile Ala Gln Leu Lys Asp Leu Leu Pro Glu
 35 40 45

His Leu Lys Leu Thr Thr Leu Gly His Leu Glu Lys Ala Val Val Leu
 50 55 60

Glu Leu Thr Leu Lys His Val Lys Ala Leu Thr Asn Leu Ile Asp Gln
 65 70 75 80

Gln Gln Gln Lys Ile Ile Ala Leu Gln Ser Gly Leu Gln Ala Gly Glu
 85 90 95

Leu Ser Gly Arg Asn Val Glu Thr Gly Gln Glu Met Phe Cys Ser Gly
 100 105 110

Phe Gln Thr Cys Ala Arg Glu Val Leu Gln Tyr Leu Ala Lys His Glu
 115 120 125

Asn Thr Arg Asp Leu Lys Ser Ser Gln Leu Val Thr His Leu His Arg
 130 135 140

Val Val Ser Glu Leu Leu Gln Gly Gly Thr Ser Arg Lys Pro Ser Asp
 145 150 155 160

Pro Ala Pro Lys Val Met Asp Phe Lys Glu Lys Pro Ser Ser Pro Ala
 165 170 175

Lys Gly Ser Glu Gly Pro Gly Lys Asn Cys Val Pro Val Ile Gln Arg
 180 185 190

Thr Phe Ala His Ser Ser Gly Glu Gln Ser Gly Ser Asp Thr Asp Thr
 195 200 205

193

Asp Ser Gly Tyr Gly Gly Glu Ser Glu Lys Gly Asp Leu Arg Ser Glu
 210 215 220

Gln Pro Cys Phe Lys Ser Asp His Gly Arg Arg Phe Thr Met Gly Glu
 225 230 235 240

Arg Ile Gly Ala Ile Lys Gln Glu Ser Glu Glu Pro Pro Thr Lys Lys
 245 250 255

Asn Arg Met Gln Leu Ser Asp Asp Glu Gly His Phe Thr Ser Ser Asp
 260 265 270

Leu Ile Ser Ser Pro Phe Leu Gly Pro His Pro His Gln Pro Pro Phe
 275 280 285

Cys Leu Pro Phe Tyr Leu Ile Pro Pro Ser Ala Thr Ala Tyr Leu Pro
 290 295 300

Met Leu Glu Lys Cys Trp Tyr Pro Thr Ser Val Pro Val Leu Tyr Pro
 305 310 315 320

Gly Leu Asn Ala Ser Ala Ala Ala Leu Ser Ser Phe Met Asn Pro Asp
 325 330 335

Lys Ile Ser Ala Pro Leu Leu Met Pro Gln Arg Leu Pro Ser Pro Leu
 340 345 350

Pro Ala His Pro Ser Val Asp Ser Ser Val Leu Leu Gln Ala Leu Lys
 355 360 365

Pro Ile Pro Pro Leu Asn Leu Glu Thr Lys Asp
 370 375

<210> 157
 <211> 358
 <212> PRT
 <213> Homo sapien

<400> 157

Met Lys Pro Gly Glu Gly Gly Gln Val Ala Pro Ser Leu Pro Gly Ser
 1 5 10 15

Gly Gln Thr Cys Leu Glu Ser Gln Gly Arg Thr Arg Ser Ser Asn Pro
 20 25 30

Pro Thr Ala Pro Ser Arg Leu Pro Ala Leu Pro His Phe Ser Phe Thr
 35 40 45

194

Trp Leu Ala Arg Arg Arg Gln Thr Ala Gln Gly Ala His Thr Ala Ser
 50 55 60

Leu Cys Ala Glu Ser Glu Pro Glu Ala Ala Gly Thr Pro Gly His Ala
 65 70 75 80

Arg Pro Gln Leu Lys Leu His Leu Lys Ala Glu Asp Ser Ser Ser Pro
 85 90 95

Gly Asp Phe Lys Glu Leu Arg Leu Arg Gly Thr Ser Ala Glu Arg Pro
 100 105 110

Pro Lys Pro Ser Pro Gly Gln Ser Ser Ser Arg Arg Ser Ala Ser Ala
 115 120 125

Asp Arg Ser Ala Gln Trp Pro Arg Leu Ala Ala Pro Trp Ser Gly Ser
 130 135 140

Pro Ala Arg Asn His Pro Pro Pro Ala Cys Pro Lys His Arg Asp Trp
 145 150 155 160

Ser Thr Glu Thr Tyr Gln Gly Lys Leu Ala Leu Leu Gly Pro Ser Ser
 165 170 175

Leu Asn Cys Ser Pro Met Leu Cys Ala Thr Leu Asn Leu Glu Gln Leu
 180 185 190

Arg Ala His Arg Glu Val Leu Ala Arg Gln Asn Ala Cys Ser Arg Ala
 195 200 205

Gln Ala Val Thr Thr Leu Pro Gly Leu Ser Ser Cys Arg Met Tyr Pro
 210 215 220

Ala His Met Tyr Gln Val Tyr Lys Ser Arg Arg Gly Ile Lys Arg Ser
 225 230 235 240

Glu Asp Ser Lys Val Ser Lys Cys Thr Pro Arg Asp Pro Ala Leu Ser
 245 250 255

Pro Ser Arg Ala Leu Ser Phe Gln Glu Lys Phe Ser Arg Phe Glu Val
 260 265 270

Gly Glu Gly Met Gln Gly Val Gly Ser Val Pro Leu Leu Ser Asp Leu
 275 280 285

195

Glu Lys Lys Gly Gln Thr Met Val Leu Gly Ala Thr Leu Leu Leu Cys
 290 295 300

Ser Ser Ala Gly Leu Leu Leu Arg Gly Trp Glu Asp Arg Leu Leu Ile
 305 310 315 320

Ser Phe Pro Lys Arg Pro Pro Pro Pro Arg Ala Ser Cys Arg Arg Pro
 325 330 335

Thr Asn Cys Arg Thr Gly Ser Ser Arg Lys Arg Asp Val Thr Gly Leu
 340 345 350

Thr Ser Ala Ser Pro Ser
 355

<210> 158
 <211> 329
 <212> PRT
 <213> Homo sapien

<400> 158

Leu Gln Pro Thr His Arg Ser Leu Pro Pro Pro Arg Pro Pro His Phe
 1 5 10 15

Ser Phe Thr Trp Leu Ala Arg Arg Arg Gln Thr Ala Gln Gly Ala His
 20 25 30

Thr Ala Ser Leu Cys Ala Glu Ser Glu Pro Glu Ala Ala Gly Thr Pro
 35 40 45

Gly His Ala Arg Pro Gln Leu Lys Leu His Leu Lys Ala Glu Asp Ser
 50 55 60

Ser Ser Pro Gly Asp Phe Lys Glu Leu Arg Leu Arg Gly Thr Ser Ala
 65 70 75 80

Glu Arg Pro Pro Lys Pro Ser Pro Gly Gln Ser Ser Ser Arg Arg Ser
 85 90 95

Ala Ser Ala Asp Arg Ser Ala Gln Trp Pro Arg Leu Ala Ala Pro Trp
 100 105 110

Ser Gly Ser Pro Ala Arg Asn His Pro Pro Pro Ala Cys Pro Lys His
 115 120 125

Arg Asp Trp Ser Thr Glu Thr Tyr Gln Gly Lys Leu Ala Leu Leu Gly

196

130

135

140

Pro Ser Ser Leu Asn Cys Ser Pro Met Leu Cys Ala Thr Leu Asn Leu
 145 150 155 160

Glu Gln Leu Arg Ala His Arg Glu Val Leu Ala Arg Gln Asn Ala Cys
 165 170 175

Ser Arg Ala Gln Ala Val Thr Thr Leu Pro Gly Leu Ser Ser Cys Arg
 180 185 190

Met Tyr Pro Ala His Met Tyr Gln Val Tyr Lys Ser Arg Arg Gly Ile
 195 200 205

Lys Arg Ser Glu Asp Ser Lys Val Ser Lys Cys Thr Pro Arg Asp Pro
 210 215 220

Ala Leu Ser Pro Ser Arg Ala Leu Ser Phe Gln Glu Lys Phe Ser Arg
 225 230 235 240

Phe Glu Val Gly Glu Gly Met Gln Gly Val Gly Ser Val Pro Leu Leu
 245 250 255

Ser Asp Leu Glu Lys Lys Gly Gln Thr Met Val Leu Gly Ala Thr Leu
 260 265 270

Leu Leu Cys Ser Ser Ala Gly Leu Leu Leu Arg Gly Trp Glu Asp Arg
 275 280 285

Leu Leu Ile Ser Phe Pro Lys Arg Pro Pro Pro Pro Arg Ala Ser Cys
 290 295 300

Arg Arg Pro Thr Asn Cys Arg Thr Gly Ser Ser Arg Lys Arg Asp Val
 305 310 315 320

Thr Gly Leu Thr Ser Ala Ser Pro Ser
 325

<210> 159

<211> 425

<212> PRT

<213> Homo sapien

<400> 159

Cys Arg Gln Glu Arg Ala Val Ala Pro Ala Arg Arg Ala Met Glu Arg
 1 5 10 15

197

Ile Pro Ser Ala Gln Pro Pro Pro Ala Cys Leu Pro Lys Ala Pro Gly
 20 25 30

Leu Glu His Gly Asp Leu Pro Gly Met Tyr Pro Ala His Met Tyr Gln
 35 40 45

Val Tyr Lys Ser Arg Arg Gly Ile Lys Arg Ser Glu Asp Ser Lys Glu
 50 55 60

Thr Tyr Lys Leu Pro His Arg Leu Ile Glu Lys Lys Arg Arg Asp Arg
 65 70 75 80

Ile Asn Glu Cys Ile Ala Gln Leu Lys Asp Leu Leu Pro Glu His Leu
 85 90 95

Lys Leu Thr Thr Leu Gly His Leu Glu Lys Ala Val Val Leu Glu Leu
 100 105 110

Thr Leu Lys His Val Lys Ala Leu Thr Asn Leu Ile Asp Gln Gln Gln
 115 120 125

Gln Lys Ile Ile Ala Leu Gln Ser Gly Leu Gln Ala Gly Glu Leu Ser
 130 135 140

Gly Arg Asn Val Glu Thr Gly Gln Glu Met Phe Cys Ser Gly Phe Gln
 145 150 155 160

Thr Cys Ala Arg Glu Val Leu Gln Tyr Leu Ala Lys His Glu Asn Thr
 165 170 175

Arg Asp Leu Lys Ser Ser Gln Leu Val Thr His Leu His Arg Val Val
 180 185 190

Ser Glu Leu Leu Gln Gly Gly Thr Ser Arg Lys Pro Ser Asp Pro Ala
 195 200 205

Pro Lys Val Met Asp Phe Lys Glu Lys Pro Ser Ser Pro Ala Lys Gly
 210 215 220

Ser Glu Gly Pro Gly Lys Asn Cys Val Pro Val Ile Gln Arg Thr Phe
 225 230 235 240

Ala His Ser Ser Gly Glu Gln Ser Gly Ser Asp Thr Asp Thr Asp Ser
 245 250 255

198

Gly Tyr Gly Gly Glu Ser Glu Lys Gly Asp Leu Arg Ser Glu Gln Pro
 260 265 270

Cys Phe Lys Ser Asp His Gly Arg Arg Phe Thr Met Gly Glu Arg Ile
 275 280 285

Gly Ala Ile Lys Gln Glu Ser Glu Glu Pro Pro Thr Lys Lys Asn Arg
 290 295 300

Met Gln Leu Ser Asp Asp Glu Gly His Phe Thr Ser Ser Asp Leu Ile
 305 310 315 320

Ser Ser Pro Phe Leu Gly Pro His Pro His Gln Pro Pro Phe Cys Leu
 325 330 335

Pro Phe Tyr Leu Ile Pro Pro Ser Ala Thr Ala Tyr Leu Pro Met Leu
 340 345 350

Glu Lys Cys Trp Tyr Pro Thr Ser Val Pro Val Leu Tyr Pro Gly Leu
 355 360 365

Asn Ala Ser Ala Ala Ala Leu Ser Ser Phe Met Asn Pro Asp Lys Ile
 370 375 380

Ser Ala Pro Leu Leu Met Pro Gln Arg Leu Pro Ser Pro Leu Pro Ala
 385 390 395 400

His Pro Ser Val Asp Ser Ser Val Leu Leu Gln Ala Leu Lys Pro Ile
 405 410 415

Pro Pro Leu Asn Leu Glu Thr Lys Asp
 420 425

<210> 160
 <211> 425
 <212> PRT
 <213> Homo sapien

<400> 160

Cys Arg Gln Glu Arg Ala Val Ala Pro Ala Arg Arg Ala Met Glu Arg
 1 5 10 15

Ile Pro Ser Ala Gln Pro Pro Pro Ala Cys Leu Pro Lys Ala Pro Gly
 20 25 30

Leu Glu His Gly Asp Leu Pro Gly Met Tyr Pro Ala His Met Tyr Gln
 35 40 45

199

Val Tyr Lys Ser Arg Arg Gly Ile Lys Arg Ser Glu Asp Ser Lys Glu
 50 55 60

Thr Tyr Lys Leu Pro His Arg Leu Ile Glu Lys Lys Arg Arg Asp Arg
 65 70 75 80

Ile Asn Glu Cys Ile Ala Gln Leu Lys Asp Leu Leu Pro Glu His Leu
 85 90 95

Lys Leu Thr Thr Leu Gly His Leu Glu Lys Ala Val Val Leu Glu Leu
 100 105 110

Thr Leu Lys His Val Lys Ala Leu Thr Asn Leu Ile Asp Gln Gln Gln
 115 120 125

Gln Lys Ile Ile Ala Leu Gln Ser Gly Leu Gln Ala Gly Glu Leu Ser
 130 135 140

Gly Arg Asn Val Glu Thr Gly Gln Glu Met Phe Cys Ser Gly Phe Gln
 145 150 155 160

Thr Cys Ala Arg Glu Val Leu Gln Tyr Leu Ala Lys His Glu Asn Thr
 165 170 175

Arg Asp Leu Lys Ser Ser Gln Leu Val Thr His Leu His Arg Val Val
 180 185 190

Ser Glu Leu Leu Gln Gly Gly Thr Ser Arg Lys Pro Ser Asp Pro Ala
 195 200 205

Pro Lys Val Met Asp Phe Lys Glu Lys Pro Ser Ser Pro Ala Lys Gly
 210 215 220

Ser Glu Gly Pro Gly Lys Asn Cys Val Pro Val Ile Gln Arg Thr Phe
 225 230 235 240

Ala His Ser Ser Gly Glu Gln Ser Gly Ser Asp Thr Asp Thr Asp Ser
 245 250 255

Gly Tyr Gly Gly Glu Ser Glu Lys Gly Asp Leu Arg Ser Glu Gln Pro
 260 265 270

Cys Phe Lys Ser Asp His Gly Arg Arg Phe Thr Met Gly Glu Arg Ile
 275 280 285

200

Gly Ala Ile Lys Gln Glu Ser Glu Glu Pro Pro Thr Lys Lys Asn Arg
 290 295 300

Met Gln Leu Ser Asp Asp Glu Gly His Phe Thr Ser Ser Asp Leu Ile
 305 310 315 320

Ser Ser Pro Phe Leu Gly Pro His Pro His Gln Pro Pro Phe Cys Leu
 325 330 335

Pro Phe Tyr Leu Ile Pro Pro Ser Ala Thr Ala Tyr Leu Pro Met Leu
 340 345 350

Glu Lys Cys Trp Tyr Pro Thr Ser Val Pro Val Leu Tyr Pro Gly Leu
 355 360 365

Asn Ala Ser Ala Ala Ala Leu Ser Ser Phe Met Asn Pro Asp Lys Ile
 370 375 380

Ser Ala Pro Leu Leu Met Pro Gln Arg Leu Pro Ser Pro Leu Pro Ala
 385 390 395 400

His Pro Ser Val Asp Ser Ser Val Leu Leu Gln Ala Leu Lys Pro Ile
 405 410 415

Pro Pro Leu Asn Leu Glu Thr Lys Asp
 420 425

<210> 161
 <211> 64
 <212> PRT
 <213> Homo sapien

<400> 161

His Val Leu Glu Leu Leu Pro Gly Gln Leu Glu Gln Asp Asp Ser Gly
 1 5 10 15

Pro Gly Val Thr Ser Gly Gln Cys Ala Gly Val Lys Asp Leu Thr Gly
 20 25 30

Leu Arg Arg Asp Leu Arg Phe Arg Pro Gly Ser Gly Ala Val Lys Leu
 35 40 45

Pro Val Glu Leu Ala Leu Ala Phe Arg Asn Ser Ser Ser Phe Cys Arg
 50 55 60

<210> 162

201

<211> 111
 <212> PRT
 <213> Homo sapien

<400> 162

Asn Phe Lys Gln Ala Val Ser Thr Gly Leu Asn Ser Pro His Pro His
 1 5 10 15

Gln Pro Pro Phe Cys Leu Pro Phe Tyr Leu Ile Pro Pro Ser Ala Thr
 20 25 30

Ala Tyr Leu Pro Met Leu Glu Lys Cys Trp Tyr Pro Thr Ser Val Pro
 35 40 45

Val Leu Tyr Pro Gly Leu Asn Ala Ser Ala Ala Ala Leu Ser Ser Phe
 50 55 60

Met Asn Pro Asp Lys Ile Ser Ala Pro Leu Leu Met Pro Gln Arg Leu
 65 70 75 80

Pro Ser Pro Leu Pro Ala His Pro Ser Val Asp Ser Ser Val Leu Leu
 85 90 95

Gln Ala Leu Lys Pro Ile Pro Pro Leu Asn Leu Glu Thr Lys Asp
 100 105 110

<210> 163
 <211> 145
 <212> PRT
 <213> Homo sapien

<400> 163

Met Gly Lys Ser Arg Cys Pro Glu Gly Phe Pro Ile Ala Glu Val Phe
 1 5 10 15

Thr Leu Lys Pro Leu Glu Phe Gly Lys Pro Asn Thr Leu Val Cys Phe
 20 25 30

Val Ser Asn Leu Phe Pro Pro Met Leu Thr Val Asn Trp Gln His His
 35 40 45

Ser Val Pro Val Glu Gly Phe Gly Pro Thr Phe Val Ser Ala Val Asp
 50 55 60

Gly Leu Ser Phe Gln Ala Phe Ser Tyr Leu Asn Phe Thr Pro Glu Pro
 65 70 75 80

202

Ser Asp Ile Phe Ser Cys Ile Val Thr His Glu Ile Asp Arg Tyr Thr
85 90 95

Ala Ile Ala Tyr Trp Val Pro Arg Asn Ala Leu Pro Ser Asp Leu Leu
100 105 110

Glu Asn Val Leu Cys Gly Val Ala Phe Gly Leu Gly Val Leu Gly Ile
115 120 125

Ile Val Gly Ile Val Leu Ile Ile Tyr Phe Arg Lys Pro Cys Ser Gly
130 135 140

Asp
145

<210> 164
<211> 270
<212> PRT
<213> Homo sapien

<400> 164

Leu Leu Pro Thr Val Trp Gln Glu Gly Met Gly His Glu Gln Asn Gln
1 5 10 15

Gly Ala Ala Leu Leu Gln Met Leu Pro Leu Leu Trp Leu Leu Pro His
20 25 30

Ser Trp Ala Val Pro Glu Ala Pro Thr Pro Met Trp Pro Asp Asp Leu
35 40 45

Gln Asn His Thr Phe Leu His Thr Val Tyr Cys Gln Asp Gly Ser Pro
50 55 60

Ser Val Gly Leu Ser Glu Ala Tyr Asp Glu Asp Gln Leu Phe Phe Phe
65 70 75 80

Asp Phe Ser Gln Asn Thr Arg Val Pro Arg Leu Pro Glu Phe Ala Asp
85 90 95

Trp Ala Gln Glu Gln Gly Asp Ala Pro Ala Ile Leu Phe Asp Lys Glu
100 105 110

Phe Cys Glu Trp Met Ile Gln Gln Ile Gly Pro Lys Leu Asp Gly Lys
115 120 125

Ile Pro Val Ser Arg Gly Phe Pro Ile Ala Glu Val Phe Thr Leu Lys
130 135 140

203

Pro Leu Glu Phe Gly Lys Pro Asn Thr Leu Val Cys Phe Val Ser Asn
 145 150 155 160

Leu Phe Pro Pro Met Leu Thr Val Asn Trp Gln His His Ser Val Pro
 165 170 175

Val Glu Gly Phe Gly Pro Thr Phe Val Ser Ala Val Asp Gly Leu Ser
 180 185 190

Phe Gln Ala Phe Ser Tyr Leu Asn Phe Thr Pro Glu Pro Ser Asp Ile
 195 200 205

Phe Ser Cys Ile Val Thr His Glu Ile Asp Arg Tyr Thr Ala Ile Ala
 210 215 220

Tyr Trp Val Pro Arg Asn Ala Leu Pro Ser Asp Leu Leu Glu Asn Val
 225 230 235 240

Leu Cys Gly Val Ala Phe Gly Leu Gly Val Leu Gly Ile Ile Val Gly
 245 250 255

Ile Val Leu Ile Ile Tyr Phe Arg Lys Pro Cys Ser Gly Asp
 260 265 270

<210> 165
 <211> 180
 <212> PRT
 <213> Homo sapien

<400> 165

His Ser Gly Leu Phe Leu Cys Leu Phe Val Ala Glu Leu Glu Pro Ala
 1 5 10 15

Ile Leu Phe Asp Lys Glu Phe Cys Glu Trp Met Ile Gln Gln Ile Gly
 20 25 30

Pro Lys Leu Asp Gly Lys Ile Pro Val Ser Arg Gly Phe Pro Ile Ala
 35 40 45

Glu Val Phe Thr Leu Lys Pro Leu Glu Phe Gly Lys Pro Asn Thr Leu
 50 55 60

Val Cys Phe Val Ser Asn Leu Phe Pro Pro Met Leu Thr Val Asn Trp
 65 70 75 80

204

Gln His His Ser Val Pro Val Glu Gly Phe Gly Pro Thr Phe Val Ser
85 90 95

Ala Val Asp Gly Leu Ser Phe Gln Ala Phe Ser Tyr Leu Asn Phe Thr
100 105 110

Pro Glu Pro Ser Asp Ile Phe Ser Cys Ile Val Thr His Glu Ile Asp
115 120 125

Arg Tyr Thr Ala Ile Ala Tyr Trp Val Pro Arg Asn Ala Leu Pro Ser
130 135 140

Asp Leu Leu Glu Asn Val Leu Cys Gly Val Ala Phe Gly Leu Gly Val
145 150 155 160

Leu Gly Ile Ile Val Gly Ile Val Leu Ile Ile Tyr Phe Arg Lys Pro
165 170 175

Cys Ser Gly Asp
180

<210> 166
<211> 796
<212> PRT
<213> Homo sapien

<400> 166

Met Ser Leu Asp Asp Asn Leu Ser Gly Thr Ser Gly Met Glu Val Asp
1 5 10 15

Asp Arg Val Ser Ala Leu Glu Gln Arg Leu Gln Leu Gln Glu Asp Glu
20 25 30

Leu Ala Val Leu Lys Ala Ala Leu Ala Asp Ala Leu Arg Arg Leu Arg
35 40 45

Ala Cys Glu Glu Gln Gly Ala Ala Leu Arg Ala Arg Gly Thr Pro Lys
50 55 60

Gly Arg Ala Pro Pro Arg Leu Gly Thr Thr Ala Ser Val Cys Gln Leu
65 70 75 80

Leu Lys Gly Leu Pro Thr Arg Thr Pro Leu Asn Gly Ser Gly Pro Pro
85 90 95

Arg Arg Val Gly Gly Tyr Ala Thr Ser Pro Ser Ser Pro Lys Lys Glu
100 105 110

205

Ala Thr Ser Gly Arg Ser Ser Val Arg Arg Tyr Leu Ser Pro Glu Arg
 115 120 125

Leu Ala Ser Val Arg Arg Glu Asp Pro Arg Ser Arg Thr Thr Ser Ser
 130 135 140

Ser Ser Asn Cys Ser Ala Lys Lys Glu Gly Lys Thr Lys Glu Val Ile
 145 150 155 160

Phe Ser Val Glu Asp Gly Ser Val Lys Met Phe Leu Arg Gly Arg Pro
 165 170 175

Val Pro Met Met Ile Pro Asp Glu Leu Ala Pro Thr Tyr Ser Leu Asp
 180 185 190

Thr Arg Ser Glu Leu Pro Ser Cys Arg Leu Lys Leu Glu Trp Val Tyr
 195 200 205

Gly Tyr Arg Gly Arg Asp Cys Arg Ala Asn Leu Tyr Leu Leu Pro Thr
 210 215 220

Gly Glu Ile Val Tyr Phe Val Ala Ser Val Ala Val Leu Tyr Ser Val
 225 230 235 240

Glu Glu Gln Arg Gln Arg His Tyr Leu Gly His Asn Asp Asp Ile Lys
 245 250 255

Cys Leu Ala Ile His Pro Asp Met Val Thr Ile Ala Thr Gly Gln Val
 260 265 270

Ala Gly Thr Thr Lys Glu Gly Lys Pro Leu Pro Pro His Val Arg Ile
 275 280 285

Trp Asp Ser Val Ser Leu Ser Thr Leu His Val Leu Gly Leu Gly Val
 290 295 300

Phe Asp Arg Ala Val Cys Cys Val Gly Phe Ser Lys Ser Asn Gly Gly
 305 310 315 320

Asn Leu Leu Cys Ala Val Asp Glu Ser Asn Asp His Met Leu Ser Val
 325 330 335

Trp Asp Trp Ala Lys Glu Thr Lys Val Val Asp Val Lys Cys Ser Asn
 340 345 350

206

Glu Ala Val Leu Val Ala Thr Phe His Pro Thr Asp Pro Thr Val Leu
 355 360 365

Ile Thr Cys Gly Lys Ser His Ile Tyr Phe Trp Thr Leu Glu Gly Gly
 370 375 380

Ser Leu Ser Lys Arg Gln Gly Leu Phe Glu Lys His Glu Lys Pro Lys
 385 390 395 400

Tyr Val Leu Cys Val Thr Phe Leu Glu Gly Gly Asp Val Val Thr Gly
 405 410 415

Asp Ser Gly Gly Asn Leu Tyr Val Trp Gly Lys Gly Gly Asn Arg Ile
 420 425 430

Thr Gln Ala Val Leu Gly Ala His Asp Gly Gly Val Phe Gly Leu Cys
 435 440 445

Ala Leu Arg Asp Gly Thr Leu Val Ser Gly Gly Gly Arg Asp Arg Arg
 450 455 460

Val Val Leu Trp Gly Ser Asp Tyr Ser Lys Leu Gln Glu Val Glu Val
 465 470 475 480

Pro Glu Asp Phe Gly Pro Val Arg Thr Val Ala Glu Gly His Gly Asp
 485 490 495

Thr Leu Tyr Val Gly Thr Thr Arg Asn Ser Ile Leu Gln Gly Ser Val
 500 505 510

His Thr Gly Phe Ser Leu Leu Val Gln Gly His Val Glu Glu Leu Trp
 515 520 525

Gly Leu Ala Thr His Pro Ser Arg Ala Gln Phe Val Thr Cys Gly Gln
 530 535 540

Asp Lys Leu Val His Leu Trp Ser Ser Asp Ser His Gln Pro Leu Trp
 545 550 555 560

Ser Arg Ile Ile Glu Asp Pro Ala Arg Ser Ala Gly Phe His Pro Ser
 565 570 575

Gly Ser Val Leu Ala Val Gly Thr Val Thr Gly Arg Trp Leu Leu Leu
 580 585 590

207

Asp Thr Glu Thr His Asp Leu Val Ala Ile His Thr Asp Gly Asn Glu
595 600 605

Gln Ile Ser Val Val Ser Phe Ser Pro Asp Gly Ala Tyr Leu Ala Val
610 615 620

Gly Ser His Asp Asn Leu Val Tyr Val Tyr Thr Val Asp Gln Gly Gly
625 630 635 640

Arg Lys Val Ser Arg Leu Gly Lys Cys Ser Gly His Ser Ser Phe Ile
645 650 655

Thr His Leu Asp Trp Ala Gln Asp Ser Ser Cys Phe Val Thr Asn Ser
660 665 670

Gly Asp Tyr Glu Ile Leu Tyr Trp Asp Pro Ala Thr Cys Lys Gln Ile
675 680 685

Thr Ser Ala Asp Ala Val Arg Asn Met Glu Trp Ala Thr Ala Thr Cys
690 695 700

Val Leu Gly Phe Gly Val Phe Gly Ile Trp Ser Glu Gly Ala Asp Gly
705 710 715 720

Thr Asp Ile Asn Ala Val Ala Arg Ser His Asp Gly Lys Leu Leu Ala
725 730 735

Ser Ala Asp Asp Phe Gly Lys Val His Leu Phe Ser Tyr Pro Cys Cys
740 745 750

Gln Pro Arg Ala Leu Ser His Lys Tyr Gly Gly His Ser Ser His Val
755 760 765

Thr Asn Val Ala Phe Leu Trp Asp Asp Ser Met Ala Leu Thr Thr Gly
770 775 780

Gly Lys Asp Thr Ser Val Leu Gln Trp Arg Val Val
785 790 795

<210> 167

<211> 627

<212> PRT

<213> Homo sapien

<400> 167

Met Phe Leu Arg Gly Arg Pro Val Pro Met Met Ile Pro Asp Glu Leu
1 5 10 15

208

Ala Pro Thr Tyr Ser Leu Asp Thr Arg Ser Glu Leu Pro Ser Cys Arg
 20 25 30

Leu Lys Leu Glu Trp Val Tyr Gly Tyr Arg Gly Arg Asp Cys Arg Ala
 35 40 45

Asn Leu Tyr Leu Leu Pro Thr Gly Glu Ile Val Tyr Phe Val Ala Ser
 50 55 60

Val Ala Val Leu Tyr Ser Val Glu Glu Gln Arg Gln Arg His Tyr Leu
 65 70 75 80

Gly His Asn Asp Asp Ile Lys Cys Leu Ala Ile His Pro Asp Met Val
 85 90 95

Thr Ile Ala Thr Gly Gln Val Ala Gly Thr Thr Lys Glu Gly Lys Pro
 100 105 110

Leu Pro Pro His Val Arg Ile Trp Asp Ser Val Ser Leu Ser Thr Leu
 115 120 125

His Val Leu Gly Leu Gly Val Phe Asp Arg Ala Val Cys Cys Val Gly
 130 135 140

Phe Ser Lys Ser Asn Gly Gly Asn Leu Leu Cys Ala Val Asp Glu Ser
 145 150 155 160

Asn Asp His Met Leu Ser Val Trp Asp Trp Ala Lys Glu Thr Lys Val
 165 170 175

Val Asp Val Lys Cys Ser Asn Glu Ala Val Leu Val Ala Thr Phe His
 180 185 190

Pro Thr Asp Pro Thr Val Leu Ile Thr Cys Gly Lys Ser His Ile Tyr
 195 200 205

Phe Trp Thr Leu Glu Gly Gly Ser Leu Ser Lys Arg Gln Gly Leu Phe
 210 215 220

Glu Lys His Glu Lys Pro Lys Tyr Val Leu Cys Val Thr Phe Leu Glu
 225 230 235 240

Gly Gly Asp Val Val Thr Gly Asp Ser Gly Gly Asn Leu Tyr Val Trp
 245 250 255

209

Gly Lys Gly Gly Asn Arg Ile Thr Gln Ala Val Leu Gly Ala His Asp
 260 265 270

Gly Gly Val Phe Gly Leu Cys Ala Leu Arg Asp Gly Thr Leu Val Ser
 275 280 285

Gly Gly Gly Arg Asp Arg Arg Val Val Leu Trp Gly Ser Asp Tyr Ser
 290 295 300

Lys Leu Gln Glu Val Glu Val Pro Glu Asp Phe Gly Pro Val Arg Thr
 305 310 315 320

Val Ala Glu Gly His Gly Asp Thr Leu Tyr Val Gly Thr Thr Arg Asn
 325 330 335

Ser Ile Leu Gln Gly Ser Val His Thr Gly Phe Ser Leu Leu Val Gln
 340 345 350

Gly His Val Glu Glu Leu Trp Gly Leu Ala Thr His Pro Ser Arg Ala
 355 360 365

Gln Phe Val Thr Cys Gly Gln Asp Lys Leu Val His Leu Trp Ser Ser
 370 375 380

Asp Ser His Gln Pro Leu Trp Ser Arg Ile Ile Glu Asp Pro Ala Arg
 385 390 395 400

Ser Ala Gly Phe His Pro Ser Gly Ser Val Leu Ala Val Gly Thr Val
 405 410 415

Thr Gly Arg Trp Leu Leu Leu Asp Thr Glu Thr His Asp Leu Val Ala
 420 425 430

Ile His Thr Asp Gly Asn Glu Gln Ile Ser Val Val Ser Phe Ser Pro
 435 440 445

Asp Gly Ala Tyr Leu Ala Val Gly Ser His Asp Asn Leu Val Tyr Val
 450 455 460

Tyr Thr Val Asp Gln Gly Gly Arg Lys Val Ser Arg Leu Gly Lys Cys
 465 470 475 480

Ser Gly His Ser Ser Phe Ile Thr His Leu Asp Trp Ala Gln Asp Ser
 485 490 495

210

Ser Cys Phe Val Thr Asn Ser Gly Asp Tyr Glu Ile Leu Tyr Trp Asp
 500 505 510

Pro Ala Thr Cys Lys Gln Ile Thr Ser Ala Asp Ala Val Arg Asn Met
 515 520 525

Glu Trp Ala Thr Ala Thr Cys Val Leu Gly Phe Gly Val Phe Gly Ile
 530 535 540

Trp Ser Glu Gly Ala Asp Gly Thr Asp Ile Asn Ala Val Ala Arg Ser
 545 550 555 560

His Asp Gly Lys Leu Leu Ala Ser Ala Asp Asp Phe Gly Lys Val His
 565 570 575

Leu Phe Ser Tyr Pro Cys Cys Gln Pro Arg Ala Leu Ser His Lys Tyr
 580 585 590

Gly Gly His Ser Ser His Val Thr Asn Val Ala Phe Leu Trp Asp Asp
 595 600 605

Ser Met Ala Leu Thr Thr Gly Gly Lys Asp Thr Ser Val Leu Gln Trp
 610 615 620

Arg Val Val
 625

<210> 168
 <211> 627
 <212> PRT
 <213> Homo sapien

<400> 168

Met Phe Leu Arg Gly Arg Pro Val Pro Met Met Ile Pro Asp Glu Leu
 1 5 10 15

Ala Pro Thr Tyr Ser Leu Asp Thr Arg Ser Glu Leu Pro Ser Cys Arg
 20 25 30

Leu Lys Leu Glu Trp Val Tyr Gly Tyr Arg Gly Arg Asp Cys Arg Ala
 35 40 45

Asn Leu Tyr Leu Leu Pro Thr Gly Glu Ile Val Tyr Phe Val Ala Ser
 50 55 60

Val Ala Val Leu Tyr Ser Val Glu Glu Gln Arg Gln Arg His Tyr Leu
 65 70 75 80

211

Gly His Asn Asp Asp Ile Lys Cys Leu Ala Ile His Pro Asp Met Val
 85 90 95

Thr Ile Ala Thr Gly Gln Val Ala Gly Thr Thr Lys Glu Gly Lys Pro
 100 105 110

Leu Pro Pro His Val Arg Ile Trp Asp Ser Val Ser Leu Ser Thr Leu
 115 120 125

His Val Leu Gly Leu Gly Val Phe Asp Arg Ala Val Cys Cys Val Gly
 130 135 140

Phe Ser Lys Ser Asn Gly Gly Asn Leu Leu Cys Ala Val Asp Glu Ser
 145 150 155 160

Asn Asp His Met Leu Ser Val Trp Asp Trp Ala Lys Glu Thr Lys Val
 165 170 175

Val Asp Val Lys Cys Ser Asn Glu Ala Val Leu Val Ala Thr Phe His
 180 185 190

Pro Thr Asp Pro Thr Val Leu Ile Thr Cys Gly Lys Ser His Ile Tyr
 195 200 205

Phe Trp Thr Leu Glu Gly Gly Ser Leu Ser Lys Arg Gln Gly Leu Phe
 210 215 220

Glu Lys His Glu Lys Pro Lys Tyr Val Leu Cys Val Thr Phe Leu Glu
 225 230 235 240

Gly Gly Asp Val Val Thr Gly Asp Ser Gly Gly Asn Leu Tyr Val Trp
 245 250 255

Gly Lys Gly Gly Asn Arg Ile Thr Gln Ala Val Leu Gly Ala His Asp
 260 265 270

Gly Gly Val Phe Gly Leu Cys Ala Leu Arg Asp Gly Thr Leu Val Ser
 275 280 285

Gly Gly Gly Arg Asp Arg Arg Val Val Leu Trp Gly Ser Asp Tyr Ser
 290 295 300

Lys Leu Gln Glu Val Glu Val Pro Glu Asp Phe Gly Pro Val Arg Thr
 305 310 315 320

212

Val Ala Glu Gly His Gly Asp Thr Leu Tyr Val Gly Thr Thr Arg Asn
 325 330 335

Ser Ile Leu Gln Gly Ser Val His Thr Gly Phe Ser Leu Leu Val Gln
 340 345 350

Gly His Val Glu Glu Leu Trp Gly Leu Ala Thr His Pro Ser Arg Ala
 355 360 365

Gln Phe Val Thr Cys Gly Gln Asp Lys Leu Val His Leu Trp Ser Ser
 370 375 380

Asp Ser His Gln Pro Leu Trp Ser Arg Ile Ile Glu Asp Pro Ala Arg
 385 390 395 400

Ser Ala Gly Phe His Pro Ser Gly Ser Val Leu Ala Val Gly Thr Val
 405 410 415

Thr Gly Arg Trp Leu Leu Leu Asp Thr Glu Thr His Asp Leu Val Ala
 420 425 430

Ile His Thr Asp Gly Asn Glu Gln Ile Ser Val Val Ser Phe Ser Pro
 435 440 445

Asp Gly Ala Tyr Leu Ala Val Gly Ser His Asp Asn Leu Val Tyr Val
 450 455 460

Tyr Thr Val Asp Gln Gly Gly Arg Lys Val Ser Arg Leu Gly Lys Cys
 465 470 475 480

Ser Gly His Ser Ser Phe Ile Thr His Leu Asp Trp Ala Gln Asp Ser
 485 490 495

Ser Cys Phe Val Thr Asn Ser Gly Asp Tyr Glu Ile Leu Tyr Trp Asp
 500 505 510

Pro Ala Thr Cys Lys Gln Ile Thr Ser Ala Asp Ala Val Arg Asn Met
 515 520 525

Glu Trp Ala Thr Ala Thr Cys Val Leu Gly Phe Gly Val Phe Gly Ile
 530 535 540

Trp Ser Glu Gly Ala Asp Gly Thr Asp Ile Asn Ala Val Ala Arg Ser
 545 550 555 560

His Asp Gly Lys Leu Leu Ala Ser Ala Asp Asp Phe Gly Lys Val His
565 570 575

Leu Phe Ser Tyr Pro Cys Cys Gln Pro Arg Ala Leu Ser His Lys Tyr
580 585 590

Gly Gly His Ser Ser His Val Thr Asn Val Ala Phe Leu Trp Asp Asp
595 600 605

Ser Met Ala Leu Thr Thr Gly Gly Lys Asp Thr Ser Val Leu Gln Trp
610 615 620

Arg Val Val
625

```
<210> 169
<211> 483
<212> PRT
<213> Homo sapien
```

<400> 169

Met Leu Glu Arg Arg Ala Leu Leu Trp Gln Arg Glu Ala Gly Pro Gly
1 5 10 15

Trp Gly Asp Arg Ala Arg Ala Gly Thr Gly Gly Ala Gly Gly Gly Cys
20 25 30

Gly Gly Ala Met Ala Glu Arg Gly Pro Ala Phe Cys Gly Leu Tyr Asp
35 40 45

Thr Ser Ser Leu Leu Arg Tyr Cys Asn Asp Asp Asn Leu Ser Gly Thr
50 55 60

Ser Gly Met Glu Val Asp Asp Arg Val Ser Ala Leu Glu Gln Arg Leu
65 70 75 80

Gln Leu Gln Glu Asp Glu Leu Ala Val Leu Lys Ala Ala Leu Ala Asp
85 90 95

Ala Leu Arg Arg Leu Arg Ala Cys Glu Glu Gln Gly Ala Ala Leu Arg
100 105 110

Ala Arg Gly Thr Pro Lys Gly Arg Ala Pro Pro Arg Leu Gly Thr Thr
115 120 125

Ala Ser Val Cys Gln Leu Leu Lys Gly Leu Pro Thr Arg Thr Pro Leu
130 135 140

214

Asn Gly Ser Gly Pro Pro Arg Arg Val Gly Gly Tyr Ala Thr Ser Pro
 145 150 155 160

Ser Ser Pro Lys Lys Glu Ala Thr Ser Gly Arg Ser Ser Val Arg Arg
 165 170 175

Tyr Leu Ser Pro Glu Arg Leu Ala Ser Val Arg Arg Glu Asp Pro Arg
 180 185 190

Ser Arg Thr Thr Ser Ser Ser Ser Asn Cys Ser Ala Lys Lys Glu Gly
 195 200 205

Lys Thr Lys Glu Val Ile Phe Ser Val Glu Asp Gly Ser Val Lys Met
 210 215 220

Phe Leu Arg Gly Arg Pro Val Pro Met Met Ile Pro Asp Glu Leu Ala
 225 230 235 240

Pro Thr Tyr Ser Leu Asp Thr Arg Ser Glu Leu Pro Ser Cys Arg Leu
 245 250 255

Lys Leu Glu Trp Val Tyr Gly Tyr Arg Gly Arg Asp Cys Arg Ala Asn
 260 265 270

Leu Tyr Leu Leu Pro Thr Gly Glu Ile Val Tyr Phe Val Ala Ser Val
 275 280 285

Ala Val Leu Tyr Ser Val Glu Glu Gln Arg Gln Arg His Tyr Leu Gly
 290 295 300

His Asn Asp Asp Ile Lys Cys Leu Ala Ile His Pro Asp Met Val Thr
 305 310 315 320

Ile Ala Thr Gly Gln Val Ala Gly Thr Thr Lys Glu Gly Lys Pro Leu
 325 330 335

Pro Pro His Val Arg Ile Trp Asp Ser Val Ser Leu Ser Thr Leu His
 340 345 350

Val Leu Gly Leu Gly Val Phe Asp Arg Ala Val Cys Cys Val Gly Phe
 355 360 365

Ser Lys Ser Asn Gly Gly Asn Leu Leu Cys Ala Val Asp Glu Ser Asn
 370 375 380

215

Asp His Met Leu Ser Val Trp Asp Trp Ala Lys Glu Thr Lys Val Val
 385 390 395 400

Asp Val Lys Cys Ser Asn Glu Ala Val Leu Val Ala Thr Phe His Pro
 405 410 415

Thr Asp Pro Thr Val Leu Ile Thr Cys Gly Lys Ser His Ile Tyr Phe
 420 425 430

Trp Thr Leu Glu Gly Gly Ser Leu Ser Lys Arg Gln Gly Leu Phe Glu
 435 440 445

Lys His Glu Lys Pro Lys Tyr Val Leu Cys Val Thr Phe Leu Glu Gly
 450 455 460

Gly Asp Val Val Thr Gly Asp Ser Gly Gly Asn Leu Tyr Val Trp Gly
 465 470 475 480

Lys Gly Pro

<210> 170
 <211> 605
 <212> PRT
 <213> Homo sapien

<400> 170

Met Ser Ser Phe Gly Ala Gly Lys Thr Lys Glu Val Ile Phe Ser Val
 1 5 10 15

Glu Asp Gly Ser Val Lys Met Phe Leu Arg Gly Arg Pro Val Pro Met
 20 25 30

Met Ile Pro Asp Glu Leu Ala Pro Thr Tyr Ser Leu Asp Thr Arg Ser
 35 40 45

Glu Leu Pro Ser Cys Arg Leu Lys Leu Glu Trp Val Tyr Gly Tyr Arg
 50 55 60

Gly Arg Asp Cys Arg Ala Asn Leu Tyr Leu Leu Pro Thr Gly Glu Ile
 65 70 75 80

Val Tyr Phe Val Ala Ser Val Ala Val Leu Tyr Ser Val Glu Glu Gln
 85 90 95

Arg Gln Arg His Tyr Leu Gly His Asn Asp Asp Ile Lys Cys Leu Ala

216

100	105	110
Ile His Pro Asp Met Val Thr 115	Ile Ala Thr Gly Gln Val 120	Ala Gly Thr 125
Thr Lys Glu Gly Lys Pro Leu Pro Pro His Val Arg Ile Trp Asp Ser 130	135	140
Val Ser Leu Ser Thr Leu His Val Leu Gly Leu Gly Val Phe Asp Arg 145	150	155 160
Ala Val Cys Cys Val Gly Phe Ser Lys Ser Asn Gly Gly Asn Leu Leu 165	170	175
Cys Ala Val Asp Glu Ser Asn Asp His Met Leu Ser Val Trp Asp Trp 180	185	190
Ala Lys Glu Thr Lys Val Val Asp Val Lys Cys Ser Asn Glu Ala Val 195	200	205
Leu Val Ala Thr Phe His Pro Thr Asp Pro Thr Val Leu Ile Thr Cys 210	215	220
Gly Lys Ser His Ile Tyr Phe Trp Thr Leu Glu Gly Gly Ser Leu Ser 225	230	235 240
Lys Arg Gln Gly Leu Phe Glu Lys His Glu Lys Pro Lys Tyr Val Leu 245	250	255
Cys Val Thr Phe Leu Glu Gly Gly Asp Val Val Thr Gly Asp Ser Gly 260	265	270
Gly Asn Leu Tyr Val Trp Gly Lys Gly Gly Asn Arg Ile Thr Gln Ala 275	280	285
Val Leu Gly Ala His Asp Gly Gly Val Phe Gly Leu Cys Ala Leu Arg 290	295	300
Asp Gly Thr Leu Val Ser Gly Gly Gly Arg Asp Arg Arg Val Val Leu 305	310	315 320
Trp Gly Ser Asp Tyr Ser Lys Leu Gln Glu Val Glu Val Pro Glu Asp 325	330	335
Phe Gly Pro Val Arg Thr Val Ala Glu Gly His Gly Asp Thr Leu Tyr 340	345	350

217

Val Gly Thr Thr Arg Asn Ser Ile Leu Gln Gly Ser Val His Thr Gly
 355 360 365

Phe Ser Leu Leu Val Gln Asp Pro Ala Arg Ser Ala Gly Phe His Pro
 370 375 380

Ser Gly Ser Val Leu Ala Val Gly Thr Val Thr Gly Arg Trp Leu Leu
 385 390 395 400

Leu Asp Thr Glu Thr His Asp Leu Val Ala Ile His Thr Asp Gly Asn
 405 410 415

Glu Gln Ile Ser Val Val Ser Phe Ser Pro Asp Gly Ala Tyr Leu Ala
 420 425 430

Val Gly Ser His Asp Asn Leu Val Tyr Val Tyr Thr Val Asp Gln Gly
 435 440 445

Gly Arg Lys Val Ser Arg Leu Gly Lys Cys Ser Gly His Ser Ser Phe
 450 455 460

Ile Thr His Leu Asp Trp Ala Gln Asp Ser Ser Cys Phe Val Thr Asn
 465 470 475 480

Ser Gly Asp Tyr Glu Ile Leu Tyr Trp Asp Pro Ala Thr Cys Lys Gln
 485 490 495

Ile Thr Ser Ala Asp Ala Val Arg Asn Met Glu Trp Ala Thr Ala Thr
 500 505 510

Cys Val Leu Gly Phe Gly Val Phe Gly Ile Trp Ser Glu Gly Ala Asp
 515 520 525

Gly Thr Asp Ile Asn Ala Val Ala Arg Ser His Asp Gly Lys Leu Leu
 530 535 540

Ala Ser Ala Asp Asp Phe Gly Lys Val His Leu Phe Ser Tyr Pro Cys
 545 550 555 560

Cys Gln Pro Arg Ala Leu Ser His Lys Tyr Gly Gly His Ser Ser His
 565 570 575

Val Thr Asn Val Ala Phe Leu Trp Asp Asp Ser Met Ala Leu Thr Thr
 580 585 590

218

Gly Gly Lys Asp Thr Ser Val Leu Gln Trp Arg Val Val
 595 600 605

<210> 171
 <211> 495
 <212> PRT
 <213> Homo sapien

<400> 171

Met Ser Ser Phe Gly Ala Gly Lys Thr Lys Glu Val Ile Phe Ser Val
 1 5 10 15

Glu Asp Gly Ser Val Lys Met Phe Leu Arg Gly Arg Pro Val Pro Met
 20 25 30

Met Ile Pro Asp Glu Leu Ala Pro Thr Tyr Ser Leu Asp Thr Arg Ser
 35 40 45

Glu Leu Pro Ser Cys Arg Leu Lys Leu Glu Trp Val Tyr Gly Tyr Arg
 50 55 60

Gly Arg Asp Cys Arg Ala Asn Leu Tyr Leu Leu Pro Thr Gly Glu Ile
 65 70 75 80

Val Tyr Phe Val Ala Ser Val Ala Val Leu Tyr Ser Val Glu Glu Gln
 85 90 95

Arg Gln Arg His Tyr Leu Gly His Asn Asp Asp Ile Lys Cys Leu Ala
 100 105 110

Ile His Pro Asp Met Val Thr Ile Ala Thr Gly Gln Val Ala Gly Thr
 115 120 125

Thr Lys Glu Gly Lys Pro Leu Pro Pro His Val Arg Ile Trp Asp Ser
 130 135 140

Val Ser Leu Ser Thr Leu His Val Leu Gly Leu Gly Val Phe Asp Arg
 145 150 155 160

Ala Val Cys Cys Val Gly Phe Ser Lys Ser Asn Gly Gly Asn Leu Leu
 165 170 175

Cys Ala Val Asp Glu Ser Asn Asp His Met Leu Ser Val Trp Asp Trp
 180 185 190

Ala Lys Glu Thr Lys Val Val Asp Val Lys Cys Ser Asn Glu Ala Val

219

195	200	205
Leu Val Ala Thr Phe His Pro Thr Asp Pro Thr Val Leu Ile Thr Cys		
210	215	220
Gly Lys Ser His Ile Tyr Phe Trp Thr Leu Glu Gly Gly Ser Leu Ser		
225	230	235
Lys Arg Gln Gly Leu Phe Glu Lys His Glu Lys Pro Lys Tyr Val Leu		
245	250	255
Cys Val Thr Phe Leu Glu Gly Gly Asp Val Val Thr Gly Asp Ser Gly		
260	265	270
Gly Asn Leu Tyr Val Trp Gly Lys Gly Gly Asn Arg Ile Thr Gln Ala		
275	280	285
Val Leu Gly Ala His Asp Gly Gly Val Phe Gly Leu Cys Ala Leu Arg		
290	295	300
Asp Gly Thr Leu Val Ser Gly Gly Gly Arg Asp Arg Arg Val Val Leu		
305	310	315
Trp Gly Ser Asp Tyr Ser Lys Leu Gln Glu Val Glu Val Pro Glu Asp		
325	330	335
Phe Gly Pro Val Arg Thr Val Ala Glu Gly His Gly Asp Thr Leu Tyr		
340	345	350
Val Gly Thr Thr Arg Asn Ser Ile Leu Gln Gly Ser Val His Thr Gly		
355	360	365
Phe Ser Leu Leu Val Gln Gly His Val Glu Glu Leu Trp Gly Leu Ala		
370	375	380
Thr His Pro Ser Arg Ala Gln Phe Val Thr Cys Gly Gln Asp Lys Leu		
385	390	395
Val His Leu Trp Ser Ser Asp Ser His Gln Pro Leu Trp Ser Arg Ile		
405	410	415
Ile Glu Asp Pro Ala Arg Ser Ala Gly Phe His Pro Ser Gly Ser Val		
420	425	430
Leu Ala Val Gly Thr Val Thr Gly Arg Trp Leu Leu Leu Asp Thr Glu		
435	440	445

220

Thr His Asp Leu Val Ala Ile His Thr Asp Gly Asn Glu Gln Ile Ser
 450 455 460

Val Val Ser Phe Ser Pro Gly Pro Phe Gln Phe Tyr His Pro Pro Gly
 465 470 475 480

Leu Gly Pro Gly Gln Gln Leu Leu Cys His Gln Leu Arg Gly Leu
 485 490 495

<210> 172

<211> 536

<212> PRT

<213> Homo sapien

<400> 172

Ile Gly Arg Gly Arg Pro Gly Gln Val Ala Gly Thr Thr Lys Glu Gly
 1 5 10 15

Lys Pro Leu Pro Pro His Val Arg Ile Trp Asp Ser Val Ser Leu Ser
 20 25 30

Thr Leu His Val Leu Gly Leu Gly Val Phe Asp Arg Ala Val Cys Cys
 35 40 45

Val Gly Phe Ser Lys Ser Cys Ser Asn Glu Ala Val Leu Val Ala Thr
 50 55 60

Phe His Pro Thr Asp Pro Thr Val Leu Ile Thr Cys Gly Lys Ser His
 65 70 75 80

Ile Tyr Phe Trp Thr Leu Glu Gly Gly Ser Leu Ser Lys Arg Gln Gly
 85 90 95

Leu Phe Glu Lys His Glu Lys Pro Lys Tyr Val Leu Cys Val Thr Phe
 100 105 110

Leu Glu Gly Gly Asp Val Val Thr Gly Asp Ser Gly Gly Asn Leu Tyr
 115 120 125

Val Trp Gly Lys Gly Gly Asn Arg Ile Thr Gln Ala Val Leu Gly Ala
 130 135 140

His Asp Gly Gly Val Phe Gly Leu Cys Ala Leu Arg Asp Gly Thr Leu
 145 150 155 160

221

Val Ser Gly Gly Gly Arg Asp Arg Arg Val Val Leu Trp Gly Ser Asp
 165 170 175

Tyr Ser Lys Leu Gln Glu Val Glu Val Pro Glu Asp Phe Gly Pro Val
 180 185 190

Arg Thr Val Ala Glu Gly His Gly Asp Thr Leu Tyr Val Gly Thr Thr
 195 200 205

Arg Asn Ser Ile Leu Gln Gly Ser Val His Thr Gly Phe Ser Leu Leu
 210 215 220

Val Gln Asp Pro Ala Thr Lys Ser Leu Thr Pro Ser Thr Ala Glu Gly
 225 230 235 240

Pro Gln Ala Pro Ala Pro Thr Val Leu Pro Pro Ala Thr Leu Ile Gly
 245 250 255

Gly Gly Thr Leu Gln Gly His Val Glu Glu Leu Trp Gly Leu Ala Thr
 260 265 270

His Pro Ser Arg Ala Gln Phe Val Thr Cys Gly Gln Asp Lys Leu Val
 275 280 285

His Leu Trp Ser Ser Asp Ser His Gln Pro Leu Trp Ser Arg Ile Ile
 290 295 300

Glu Asp Pro Ala Arg Ser Ala Gly Phe His Pro Ser Gly Ser Val Leu
 305 310 315 320

Ala Val Gly Thr Val Thr Gly Arg Trp Leu Leu Leu Asp Thr Glu Thr
 325 330 335

His Asp Leu Val Ala Ile His Thr Asp Gly Asn Glu Gln Ile Ser Val
 340 345 350

Val Ser Phe Ser Pro Asp Gly Ala Tyr Leu Ala Val Gly Ser His Asp
 355 360 365

Asn Leu Val Tyr Val Tyr Thr Val Asp Gln Gly Gly Arg Lys Val Ser
 370 375 380

Arg Leu Gly Lys Cys Ser Gly His Ser Ser Phe Ile Thr His Leu Asp
 385 390 395 400

Trp Ala Gln Asp Ser Ser Cys Phe Val Thr Asn Ser Gly Asp Tyr Glu

222

405

410

415

Ile Leu Tyr Trp Asp Pro Ala Thr Cys Lys Gln Ile Thr Ser Ala Asp
 420 425 430

Ala Val Arg Asn Met Glu Trp Ala Thr Ala Thr Cys Val Leu Gly Phe
 435 440 445

Gly Val Phe Gly Ile Trp Ser Glu Gly Ala Asp Gly Thr Asp Ile Asn
 450 455 460

Ala Val Ala Arg Ser His Asp Gly Lys Leu Leu Ala Ser Ala Asp Asp
 465 470 475 480

Phe Gly Lys Val His Leu Phe Ser Tyr Pro Cys Cys Gln Pro Arg Ala
 485 490 495

Leu Ser His Lys Tyr Gly Gly His Ser Ser His Val Thr Asn Val Ala
 500 505 510

Phe Leu Trp Asp Asp Ser Met Ala Leu Thr Thr Gly Gly Lys Asp Thr
 515 520 525

Ser Val Leu Gln Trp Arg Val Val
 530 535

<210> 173
 <211> 544
 <212> PRT
 <213> Homo sapien

<400> 173

Arg Leu Gly Ser Gly Leu Gly Val Asn Gly Arg Gly Arg Pro Gly Gln
 1 5 10 15

Val Ala Gly Thr Thr Lys Glu Gly Lys Pro Leu Pro Pro His Val Arg
 20 25 30

Ile Trp Asp Ser Val Ser Leu Ser Thr Leu His Val Leu Gly Leu Gly
 35 40 45

Val Phe Asp Arg Ala Val Cys Cys Val Gly Phe Ser Lys Ser Cys Ser
 50 55 60

Asn Glu Ala Val Leu Val Ala Thr Phe His Pro Thr Asp Pro Thr Val
 65 70 75 80

223

Leu Ile Thr Cys Gly Lys Ser His Ile Tyr Phe Trp Thr Leu Glu Gly
 85 90 95

Gly Ser Leu Ser Lys Arg Gln Gly Leu Phe Glu Lys His Glu Lys Pro
 100 105 110

Lys Tyr Val Leu Cys Val Thr Phe Leu Glu Gly Gly Asp Val Val Thr
 115 120 125

Gly Asp Ser Gly Gly Asn Leu Tyr Val Trp Gly Lys Gly Gly Asn Arg
 130 135 140

Ile Thr Gln Ala Val Leu Gly Ala His Asp Gly Gly Val Phe Gly Leu
 145 150 155 160

Cys Ala Leu Arg Asp Gly Thr Leu Val Ser Gly Gly Gly Arg Asp Arg
 165 170 175

Arg Val Val Leu Trp Gly Ser Asp Tyr Ser Lys Leu Gln Glu Val Glu
 180 185 190

Val Pro Glu Asp Phe Gly Pro Val Arg Thr Val Ala Glu Gly His Gly
 195 200 205

Asp Thr Leu Tyr Val Gly Thr Thr Arg Asn Ser Ile Leu Gln Gly Ser
 210 215 220

Val His Thr Gly Phe Ser Leu Leu Val Gln Asp Pro Ala Thr Lys Ser
 225 230 235 240

Leu Thr Pro Ser Thr Ala Glu Gly Pro Gln Ala Pro Ala Pro Thr Val
 245 250 255

Leu Pro Pro Ala Thr Leu Ile Gly Gly Gly Thr Leu Gln Gly His Val
 260 265 270

Glu Glu Leu Trp Gly Leu Ala Thr His Pro Ser Arg Ala Gln Phe Val
 275 280 285

Thr Cys Gly Gln Asp Lys Leu Val His Leu Trp Ser Ser Asp Ser His
 290 295 300

Gln Pro Leu Trp Ser Arg Ile Ile Glu Asp Pro Ala Arg Ser Ala Gly
 305 310 315 320

224

Phe His Pro Ser Gly Ser Val Leu Ala Val Gly Thr Val Thr Gly Arg
 325 330 335

Trp Leu Leu Leu Asp Thr Glu Thr His Asp Leu Val Ala Ile His Thr
 340 345 350

Asp Gly Asn Glu Gln Ile Ser Val Val Ser Phe Ser Pro Asp Gly Ala
 355 360 365

Tyr Leu Ala Val Gly Ser His Asp Asn Leu Val Tyr Val Tyr Thr Val
 370 375 380

Asp Gln Gly Gly Arg Lys Val Ser Arg Leu Gly Lys Cys Ser Gly His
 385 390 395 400

Ser Ser Phe Ile Thr His Leu Asp Trp Ala Gln Asp Ser Ser Cys Phe
 405 410 415

Val Thr Asn Ser Gly Asp Tyr Glu Ile Leu Tyr Trp Asp Pro Ala Thr
 420 425 430

Cys Lys Gln Ile Thr Ser Ala Asp Ala Val Arg Asn Met Glu Trp Ala
 435 440 445

Thr Ala Thr Cys Val Leu Gly Phe Gly Val Phe Gly Ile Trp Ser Glu
 450 455 460

Gly Ala Asp Gly Thr Asp Ile Asn Ala Val Ala Arg Ser His Asp Gly
 465 470 475 480

Lys Leu Leu Ala Ser Ala Asp Asp Phe Gly Lys Val His Leu Phe Ser
 485 490 495

Tyr Pro Cys Cys Gln Pro Arg Ala Leu Ser His Lys Tyr Gly Gly His
 500 505 510

Ser Ser His Val Thr Asn Val Ala Phe Leu Trp Asp Asp Ser Met Ala
 515 520 525

Leu Thr Thr Gly Gly Lys Asp Thr Ser Val Leu Gln Trp Arg Val Val
 530 535 540

<210> 174

<211> 482

<212> PRT

<213> Homo sapien

225

<220>
 <221> MISC_FEATURE
 <222> (2)..(2)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (6)..(6)
 <223> X=any amino acid

<400> 174

Ser Xaa Gly His Cys Xaa Asp Phe Ile Trp Pro Gly His Trp Leu Ser
 1 5 10 15

Thr Trp His Trp Ser Arg Gln Arg Pro Ser Trp Gly Lys Leu Met Phe
 20 25 30

Thr Gly Gly Arg Asn Pro Pro Tyr Leu Gln Ala Ala Ser Gln Pro Gln
 35 40 45

Glu Ala Thr Arg Leu Ala Glu Ser His Val Glu Ser Ala Ser Asn Met
 50 55 60

Glu Gln Leu Thr Arg Glu Thr Glu Asp Tyr Ser Lys Gln Ala Leu Ser
 65 70 75 80

Leu Val Arg Lys Ala Leu His Glu Gly Val Gly Ser Gly Ser Gly Ser
 85 90 95

Pro Asp Gly Ala Val Val Gln Gly Leu Val Glu Lys Leu Glu Lys Thr
 100 105 110

Lys Ser Leu Ala Gln Gln Leu Thr Arg Glu Ala Thr Gln Ala Glu Ile
 115 120 125

Glu Ala Asp Arg Ser Tyr Gln His Ser Leu Arg Leu Leu Asp Ser Val
 130 135 140

Ser Arg Leu Gln Gly Val Ser Asp Gln Ser Phe Gln Val Glu Glu Ala
 145 150 155 160

Lys Arg Ile Lys Gln Lys Ala Asp Ser Leu Ser Ser Leu Val Thr Arg
 165 170 175

His Met Asp Glu Phe Lys Arg Thr Gln Lys Asn Leu Gly Asn Trp Lys
 180 185 190

226

Glu Glu Ala Gln Gln Leu Leu Gln Asn Gly Lys Ser Gly Arg Glu Lys
 195 200 205

Ser Asp Gln Leu Leu Ser Arg Ala Asn Leu Ala Lys Ser Arg Ala Gln
 210 215 220

Glu Ala Leu Ser Met Gly Asn Ala Thr Phe Tyr Glu Val Glu Ser Ile
 225 230 235 240

Leu Lys Asn Leu Arg Glu Phe Asp Leu Gln Val Asp Asn Arg Lys Ala
 245 250 255

Glu Ala Glu Glu Ala Met Lys Arg Leu Ser Tyr Ile Ser Gln Lys Val
 260 265 270

Ser Asp Ala Ser Asp Lys Thr Gln Gln Ala Glu Arg Ala Leu Gly Ser
 275 280 285

Ala Ala Ala Asp Ala Gln Arg Ala Lys Asn Gly Ala Gly Glu Ala Leu
 290 295 300

Glu Ile Ser Ser Glu Ile Glu Gln Glu Ile Gly Ser Leu Asn Leu Glu
 305 310 315 320

Ala Asn Val Thr Ala Asp Gly Ala Leu Ala Met Glu Lys Gly Leu Ala
 325 330 335

Ser Leu Lys Ser Glu Met Arg Glu Val Glu Gly Glu Leu Glu Arg Lys
 340 345 350

Glu Leu Glu Phe Asp Thr Asn Met Asp Ala Val Gln Met Val Ile Thr
 355 360 365

Glu Ala Gln Lys Val Asp Thr Arg Ala Lys Asn Ala Gly Val Thr Ile
 370 375 380

Gln Asp Thr Leu Asn Thr Leu Asp Gly Leu Leu His Leu Met Asp Gln
 385 390 395 400

Pro Leu Ser Val Asp Glu Glu Gly Leu Val Leu Leu Glu Gln Lys Leu
 405 410 415

Ser Arg Ala Lys Thr Gln Ile Asn Ser Gln Leu Arg Pro Met Met Ser
 420 425 430

227

Glu Leu Glu Glu Arg Ala Arg Gln Gln Arg Gly His Leu His Leu Leu
 435 440 445

Glu Thr Ser Ile Asp Gly Ile Leu Ala Asp Val Lys Asn Leu Glu Asn
 450 455 460

Ile Arg Asp Asn Leu Pro Pro Gly Cys Tyr Asn Thr Gln Ala Leu Glu
 465 470 475 480

Gln Gln

<210> 175

<211> 454

<212> PRT

<213> Homo sapien

<400> 175

Met Leu Met Phe Thr Gly Gly Arg Asn Pro Pro Tyr Leu Gln Ala Ala
 1 5 10 15

Ser Gln Pro Gln Glu Ala Thr Arg Leu Ala Glu Ser His Val Glu Ser
 20 25 30

Ala Ser Asn Met Glu Gln Leu Thr Arg Glu Thr Glu Asp Tyr Ser Lys
 35 40 45

Gln Ala Leu Ser Leu Val Arg Lys Ala Leu His Glu Gly Val Gly Ser
 50 55 60

Gly Ser Gly Ser Pro Asp Gly Ala Val Val Gln Gly Leu Val Glu Lys
 65 70 75 80

Leu Glu Lys Thr Lys Ser Leu Ala Gln Gln Leu Thr Arg Glu Ala Thr
 85 90 95

Gln Ala Glu Ile Glu Ala Asp Arg Ser Tyr Gln His Ser Leu Arg Leu
 100 105 110

Leu Asp Ser Val Ser Arg Leu Gln Gly Val Ser Asp Gln Ser Phe Gln
 115 120 125

Val Glu Glu Ala Lys Arg Ile Lys Gln Lys Ala Asp Ser Leu Ser Ser
 130 135 140

Leu Val Thr Arg His Met Asp Glu Phe Lys Arg Thr Gln Lys Asn Leu
 145 150 155 160

228

Gly Asn Trp Lys Glu Glu Ala Gln Gln Leu Leu Gln Asn Gly Lys Ser
 165 170 175

Gly Arg Glu Lys Ser Asp Gln Leu Leu Ser Arg Ala Asn Leu Ala Lys
 180 185 190

Ser Arg Ala Gln Glu Ala Leu Ser Met Gly Asn Ala Thr Phe Tyr Glu
 195 200 205

Val Glu Ser Ile Leu Lys Asn Leu Arg Glu Phe Asp Leu Gln Val Asp
 210 215 220

Asn Arg Lys Ala Glu Ala Glu Glu Ala Met Lys Arg Leu Ser Tyr Ile
 225 230 235 240

Ser Gln Lys Val Ser Asp Ala Ser Asp Lys Thr Gln Gln Ala Glu Arg
 245 250 255

Ala Leu Gly Ser Ala Ala Ala Asp Ala Gln Arg Ala Lys Asn Gly Ala
 260 265 270

Gly Glu Ala Leu Glu Ile Ser Ser Glu Ile Glu Gln Glu Ile Gly Ser
 275 280 285

Leu Asn Leu Glu Ala Asn Val Thr Ala Asp Gly Ala Leu Ala Met Glu
 290 295 300

Lys Gly Leu Ala Ser Leu Lys Ser Glu Met Arg Glu Val Glu Gly Glu
 305 310 315 320

Leu Glu Arg Lys Glu Leu Glu Phe Asp Thr Asn Met Asp Ala Val Gln
 325 330 335

Met Val Ile Thr Glu Ala Gln Lys Val Asp Thr Arg Ala Lys Asn Ala
 340 345 350

Gly Val Thr Ile Gln Asp Thr Leu Asn Thr Leu Asp Gly Leu Leu His
 355 360 365

Leu Met Asp Gln Pro Leu Ser Val Asp Glu Glu Gly Leu Val Leu Leu
 370 375 380

Glu Gln Lys Leu Ser Arg Ala Lys Thr Gln Ile Asn Ser Gln Leu Arg
 385 390 395 400

229

Pro Met Met Ser Glu Leu Glu Glu Arg Ala Arg Gln Gln Arg Gly His
 405 410 415

Leu His Leu Leu Glu Thr Ser Ile Asp Gly Ile Leu Ala Asp Val Lys
 420 425 430

Asn Leu Glu Asn Ile Arg Asp Asn Leu Pro Pro Gly Cys Tyr Asn Thr
 435 440 445

Gln Ala Leu Glu Gln Gln
 450

<210> 176
 <211> 340
 <212> PRT
 <213> Homo sapien

<400> 176

Met His Asp Val Lys Asn His Arg Thr Phe Leu Lys Arg Thr Lys Tyr
 1 5 10 15

Asp Asn Leu His Leu Glu Asp Leu Phe Ile Gly Asn Lys Val Asn Val
 20 25 30

Phe Ser Arg Gln Leu Val Leu Ile Asp Tyr Gly Asp Gln Tyr Thr Ala
 35 40 45

Arg Gln Leu Gly Ser Arg Lys Glu Lys Thr Leu Ala Leu Ile Lys Pro
 50 55 60

Asp Ala Ile Ser Lys Ala Gly Glu Ile Ile Glu Ile Ile Asn Lys Ala
 65 70 75 80

Gly Phe Thr Ile Thr Lys Leu Lys Met Met Met Leu Ser Arg Lys Glu
 85 90 95

Ala Leu Asp Phe His Val Asp His Gln Ser Arg Pro Phe Phe Asn Glu
 100 105 110

Leu Ile Gln Phe Ile Thr Thr Gly Pro Ile Ile Ala Met Glu Ile Leu
 115 120 125

Arg Asp Asp Ala Ile Cys Glu Trp Lys Arg Leu Leu Gly Pro Ala Asn
 130 135 140

Ser Gly Val Ala Arg Thr Asp Ala Ser Glu Ser Ile Arg Ala Leu Phe

```
<210> 177
<211> 304
<212> PRT
<213> Homo sapien

<220>
<221> MISC_FEATURE
<222> (264)..(264)
<223> X=any amino acid
```

231

<400> 177

Thr Gly Pro Val Ala Met Gly Arg Val Ile Arg Gly Gln Arg Lys Gly
 1 5 10 15

Ala Gly Ser Val Phe Arg Ala His Val Lys His Arg Lys Gly Ala Ala
 20 25 30

Arg Leu Arg Ala Val Asp Phe Ala Glu Arg His Gly Tyr Ile Lys Gly
 35 40 45

Ile Val Lys Asp Ile Ile His Asp Pro Gly Arg Gly Ala Pro Leu Ala
 50 55 60

Lys Val Val Phe Arg Asp Pro Tyr Arg Phe Lys Lys Arg Thr Glu Leu
 65 70 75 80

Phe Ile Ala Ala Glu Gly Ile His Thr Gly Gln Phe Val Tyr Cys Gly
 85 90 95

Lys Lys Ala Gln Leu Asn Ile Gly Asn Val Leu Pro Val Gly Thr Met
 100 105 110

Pro Glu Gly Thr Ile Val Cys Cys Leu Glu Glu Lys Pro Gly Asp Arg
 115 120 125

Gly Lys Leu Ala Arg Ala Ser Gly Asn Tyr Ala Thr Val Ile Ser His
 130 135 140

Asn Pro Glu Thr Lys Lys Thr Arg Val Lys Leu Pro Ser Gly Ser Lys
 145 150 155 160

Lys Val Ile Ser Ser Ala Asn Arg Ala Val Val Gly Val Val Ala Gly
 165 170 175

Gly Gly Arg Ile Asp Lys Pro Ile Leu Lys Ala Gly Arg Ala Tyr His
 180 185 190

Lys Tyr Lys Ala Lys Arg Asn Cys Trp Pro Arg Val Arg Gly Val Ala
 195 200 205

Met Asn Pro Val Glu His Pro Phe Gly Gly Gly Asn His Gln His Ile
 210 215 220

Gly Lys Pro Ser Thr Ile Arg Arg Asp Ala Pro Ala Gly Arg Lys Val
 225 230 235 240

232

Gly Leu Ile Ala Ala Arg Arg Thr Gly Arg Leu Arg Gly Thr Lys Thr
 245 250 255

Val Gln Glu Asn Met Trp Ala Xaa Ser Gly Phe Ala Glu Lys Asn Thr
 260 265 270

Thr Thr Gln Thr Gln Arg Gln Thr Tyr Arg Lys Lys Gly Gly Tyr Arg
 275 280 285

Gly Arg His Ser Arg Gly Asn Ile Ile Ala Ala Glu Asp Arg Gly Gly
 290 295 300

<210> 178
 <211> 185
 <212> PRT
 <213> Homo sapien

<400> 178

Met Pro Glu Gly Thr Ile Val Cys Cys Leu Glu Glu Lys Pro Gly Asp
 1 5 10 15

Arg Gly Lys Leu Ala Arg Ala Ser Gly Asn Tyr Ala Thr Val Ile Ser
 20 25 30

His Asn Pro Glu Thr Lys Lys Thr Arg Val Lys Leu Pro Ser Gly Ser
 35 40 45

Lys Lys Val Ile Ser Ser Ala Asn Arg Ala Val Val Gly Val Val Ala
 50 55 60

Gly Gly Gly Arg Ile Asp Lys Pro Ile Leu Lys Ala Gly Arg Ala Tyr
 65 70 75 80

His Lys Tyr Lys Ala Lys Arg Asn Cys Trp Pro Arg Val Arg Gly Val
 85 90 95

Ala Met Asn Pro Val Glu His Pro Phe Gly Gly Gly Asn His Gln His
 100 105 110

Ile Gly Lys Pro Ser Thr Ile Arg Arg Asp Ala Pro Ala Gly Arg Lys
 115 120 125

Val Gly Leu Ile Ala Ala Arg Arg Thr Gly Arg Leu Arg Gly Thr Lys
 130 135 140

233

Thr Val Gln Glu Asn Met Trp Ala His Lys Trp Val Cys Arg Glu Lys
 145 150 155 160

Thr Gln Gln Arg Lys His Lys Gly Lys His Ile Glu Lys Lys Gly Ala
 165 170 175

Thr Gly Ala Asp Thr Leu Glu Val Ile
 180 185

<210> 179
 <211> 484
 <212> PRT
 <213> Homo sapien

<400> 179

His Gly Lys Arg Gly Arg His Gly Lys Arg Gly Arg His Gly Met Val
 1 5 10 15

Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala
 20 25 30

Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala
 35 40 45

Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val
 50 55 60

Ser Ala Asp Ala Met His Thr Asp Pro Asp Tyr Ser Ala Ala Tyr Val
 65 70 75 80

Val Ile Glu Thr Asp Ala Glu Asp Gly Ile Lys Gly Cys Gly Ile Thr
 85 90 95

Phe Thr Leu Gly Lys Gly Thr Glu Val Val Val Cys Ala Val Asn Ala
 100 105 110

Leu Ala His His Val Leu Asn Lys Asp Leu Lys Asp Ile Val Gly Asp
 115 120 125

Phe Arg Gly Phe Tyr Arg Gln Leu Thr Ser Asp Gly Gln Leu Arg Trp
 130 135 140

Ile Gly Pro Glu Lys Gly Val Val His Leu Ala Thr Ala Ala Val Leu
 145 150 155 160

Asn Ala Val Trp Asp Leu Trp Ala Lys Gln Glu Gly Lys Pro Val Trp
 165 170 175

234

Lys Leu Leu Val Asp Met Asp Pro Arg Met Leu Val Ser Cys Ile Asp
 180 185 190

Phe Arg Tyr Ile Thr Asp Val Leu Thr Glu Glu Asp Ala Leu Glu Ile
 195 200 205

Leu Gln Lys Gly Gln Ile Gly Lys Lys Glu Arg Glu Lys Gln Met Leu
 210 215 220

Ala Gln Gly Tyr Pro Ala Tyr Thr Thr Ser Cys Ala Trp Leu Gly Tyr
 225 230 235 240

Ser Asp Asp Thr Leu Lys Gln Leu Cys Ala Gln Ala Leu Lys Asp Gly
 245 250 255

Trp Thr Arg Phe Lys Val Lys Val Gly Ala Asp Leu Gln Asp Asp Met
 260 265 270

Arg Arg Cys Gln Ile Ile Arg Asp Met Ile Gly Pro Glu Lys Thr Leu
 275 280 285

Met Met Asp Ala Asn Gln Arg Trp Asp Val Pro Glu Ala Val Glu Trp
 290 295 300

Met Ser Lys Leu Ala Lys Phe Lys Pro Leu Trp Ile Glu Glu Pro Thr
 305 310 315 320

Ser Pro Asp Asp Ile Leu Gly His Ala Thr Ile Ser Lys Ala Leu Val
 325 330 335

Pro Leu Gly Ile Gly Ile Ala Thr Gly Glu Gln Cys His Asn Arg Val
 340 345 350

Ile Phe Lys Gln Leu Leu Gln Ala Lys Ala Leu Gln Phe Leu Gln Ile
 355 360 365

Asp Ser Cys Arg Leu Gly Ser Val Asn Glu Asn Leu Ser Val Leu Leu
 370 375 380

Met Ala Lys Lys Phe Glu Ile Pro Val Cys Pro His Ala Gly Gly Val
 385 390 395 400

Gly Leu Cys Glu Leu Val Gln His Leu Ile Ile Phe Asp Tyr Ile Ser
 405 410 415

235

Val Ser Ala Ser Leu Glu Asn Arg Val Cys Glu Tyr Val Asp His Leu
 420 425 430

His Glu His Phe Lys Tyr Pro Val Met Ile Gln Arg Ala Ser Tyr Met
 435 440 445

Pro Pro Lys Asp Pro Gly Tyr Ser Thr Glu Met Lys Glu Glu Ser Val
 450 455 460

Lys Lys His Gln Tyr Pro Asp Gly Glu Val Trp Lys Lys Leu Leu Pro
 465 470 475 480

Ala Gln Glu Asn

<210> 180
 <211> 483
 <212> PRT
 <213> Homo sapien

<400> 180

Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Trp Ser
 1 5 10 15

Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met
 20 25 30

Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp
 35 40 45

Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser
 50 55 60

Ala Asp Ala Met His Thr Asp Pro Asp Tyr Ser Ala Ala Tyr Val Val
 65 70 75 80

Ile Glu Thr Asp Ala Glu Asp Gly Ile Lys Gly Cys Gly Ile Thr Phe
 85 90 95

Thr Leu Gly Lys Gly Thr Glu Val Val Val Cys Ala Val Asn Ala Leu
 100 105 110

Ala His His Val Leu Asn Lys Asp Leu Lys Asp Ile Val Gly Asp Phe
 115 120 125

Arg Gly Phe Tyr Arg Gln Leu Thr Ser Asp Gly Gln Leu Arg Trp Ile

236

130		135		140
Gly Pro Glu Lys Gly Val Val His Leu Ala Thr Ala Ala Val Leu Asn				
145		150		155 160
Ala Val Trp Asp Leu Trp Ala Lys Gln Glu Gly Lys Pro Val Trp Lys				
	165		170	175
Leu Leu Val Asp Met Asp Pro Arg Met Leu Val Ser Cys Ile Asp Phe				
	180		185	190
Arg Tyr Ile Thr Asp Val Leu Thr Glu Glu Asp Ala Leu Glu Ile Leu				
	195		200	205
Gln Lys Gly Gln Ile Gly Lys Lys Glu Arg Glu Lys Gln Met Leu Ala				
	210		215	220
Gln Gly Tyr Pro Ala Tyr Thr Thr Ser Cys Ala Trp Leu Gly Tyr Ser				
225		230		235 240
Asp Asp Thr Leu Lys Gln Leu Cys Ala Gln Ala Leu Lys Asp Gly Trp				
	245		250	255
Thr Arg Phe Lys Val Lys Val Gly Ala Asp Leu Gln Asp Asp Met Arg				
	260		265	270
Arg Cys Gln Ile Ile Arg Asp Met Ile Gly Pro Glu Lys Thr Leu Met				
	275		280	285
Met Asp Ala Asn Gln Arg Trp Asp Val Pro Glu Ala Val Glu Trp Met				
	290		295	300
Ser Lys Leu Ala Lys Phe Lys Pro Leu Trp Ile Glu Glu Pro Thr Ser				
305		310		315 320
Pro Asp Asp Ile Leu Gly His Ala Thr Ile Ser Lys Ala Leu Val Pro				
	325		330	335
Leu Gly Ile Gly Ile Ala Thr Gly Glu Gln Cys His Asn Arg Val Ile				
	340		345	350
Phe Lys Gln Leu Leu Gln Ala Lys Ala Leu Gln Phe Leu Gln Ile Asp				
	355		360	365
Ser Cys Arg Leu Gly Ser Val Asn Glu Asn Leu Ser Val Leu Leu Met				
	370		375	380

237

Ala Lys Lys Phe Glu Ile Pro Val Cys Pro His Ala Gly Gly Val Gly
385 390 395 400

Leu Cys Glu Leu Val Gln His Leu Ile Ile Phe Asp Tyr Ile Ser Val
405 410 415

Ser Ala Ser Leu Glu Asn Arg Val Cys Glu Tyr Val Asp His Leu His
420 425 430

Glu His Phe Lys Tyr Pro Val Met Ile Gln Arg Ala Ser Tyr Met Pro
435 440 445

Pro Lys Asp Pro Gly Tyr Ser Thr Glu Met Lys Glu Glu Ser Val Lys
450 455 460

Lys His Gln Tyr Pro Asp Gly Glu Val Trp Lys Lys Leu Leu Pro Ala
465 470 475 480

Gln Glu Asn

<210> 181
<211> 484
<212> PRT
<213> Homo sapien

<400> 181

His Gly Lys Arg Gly Arg His Gly Lys Arg Gly Arg His Gly Met Val
1 5 10 15

Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala
20 25 30

Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala
35 40 45

Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val
50 55 60

Ser Ala Asp Ala Met His Thr Asp Pro Asp Tyr Ser Ala Ala Tyr Val
65 70 75 80

Val Ile Glu Thr Asp Ala Glu Asp Gly Ile Lys Gly Cys Gly Ile Thr
85 90 95

238

Phe Thr Leu Gly Lys Gly Thr Glu Val Val Val Cys Ala Val Asn Ala
 100 105 110

Leu Ala His His Val Leu Asn Lys Asp Leu Lys Asp Ile Val Gly Asp
 115 120 125

Phe Arg Gly Phe Tyr Arg Gln Leu Thr Ser Asp Gly Gln Leu Arg Trp
 130 135 140

Ile Gly Pro Glu Lys Gly Val Val His Leu Ala Thr Ala Ala Val Leu
 145 150 155 160

Asn Ala Val Trp Asp Leu Trp Ala Lys Gln Glu Gly Lys Pro Val Trp
 165 170 175

Lys Leu Leu Val Asp Met Asp Pro Arg Met Leu Val Ser Cys Ile Asp
 180 185 190

Phe Arg Tyr Ile Thr Asp Val Leu Thr Glu Glu Asp Ala Leu Glu Ile
 195 200 205

Leu Gln Lys Gly Gln Ile Gly Lys Lys Glu Arg Glu Lys Gln Met Leu
 210 215 220

Ala Gln Gly Tyr Pro Ala Tyr Thr Thr Ser Cys Ala Trp Leu Gly Tyr
 225 230 235 240

Ser Asp Asp Thr Leu Lys Gln Leu Cys Ala Gln Ala Leu Lys Asp Gly
 245 250 255

Trp Thr Arg Phe Lys Val Lys Val Gly Ala Asp Leu Gln Asp Asp Met
 260 265 270

Arg Arg Cys Gln Ile Ile Arg Asp Met Ile Gly Pro Glu Lys Thr Leu
 275 280 285

Met Met Asp Ala Asn Gln Arg Trp Asp Val Pro Glu Ala Val Glu Trp
 290 295 300

Met Ser Lys Leu Ala Lys Phe Lys Pro Leu Trp Ile Glu Glu Pro Thr
 305 310 315 320

Ser Pro Asp Asp Ile Leu Gly His Ala Thr Ile Ser Lys Ala Leu Val
 325 330 335

Pro Leu Gly Ile Gly Ile Ala Thr Gly Glu Gln Cys His Asn Arg Val

239

340

345

350

Ile Phe Lys Gln Leu Leu Gln Ala Lys Ala Leu Gln Phe Leu Gln Ile
 355 360 365

Asp Ser Cys Arg Leu Gly Ser Val Asn Glu Asn Leu Ser Val Leu Leu
 370 375 380

Met Ala Lys Lys Phe Glu Ile Pro Val Cys Pro His Ala Gly Gly Val
 385 390 395 400

Gly Leu Cys Glu Leu Val Gln His Leu Ile Ile Phe Asp Tyr Ile Ser
 405 410 415

Val Ser Ala Ser Leu Glu Asn Arg Val Cys Glu Tyr Val Asp His Leu
 420 425 430

His Glu His Phe Lys Tyr Pro Val Met Ile Gln Arg Ala Ser Tyr Met
 435 440 445

Pro Pro Lys Asp Pro Gly Tyr Ser Thr Glu Met Lys Glu Glu Ser Val
 450 455 460

Lys Lys His Gln Tyr Pro Asp Gly Glu Val Trp Lys Lys Leu Leu Pro
 465 470 475 480

Ala Gln Glu Asn

<210> 182
 <211> 484
 <212> PRT
 <213> Homo sapien

<400> 182

His Gly Lys Arg Gly Arg His Gly Lys Arg Gly Arg His Gly Met Val
 1 5 10 15

Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala
 20 25 30

Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala
 35 40 45

Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val
 50 55 60

240

Ser Ala Asp Ala Met His Thr Asp Pro Asp Tyr Ser Ala Ala Tyr Val
65 70 75 80

Val Ile Glu Thr Asp Ala Glu Asp Gly Ile Lys Gly Cys Gly Ile Thr
85 90 95

Phe Thr Leu Gly Lys Gly Thr Glu Val Val Val Cys Ala Val Asn Ala
100 105 110

Leu Ala His His Val Leu Asn Lys Asp Leu Lys Asp Ile Val Gly Asp
115 120 125

Phe Arg Gly Phe Tyr Arg Gln Leu Thr Ser Asp Gly Gln Leu Arg Trp
130 135 140

Ile Gly Pro Glu Lys Gly Val Val His Leu Ala Thr Ala Ala Val Leu
145 150 155 160

Asn Ala Val Trp Asp Leu Trp Ala Lys Gln Glu Gly Lys Pro Val Trp
165 170 175

Lys Leu Leu Val Asp Met Asp Pro Arg Met Leu Val Ser Cys Ile Asp
180 185 190

Phe Arg Tyr Ile Thr Asp Val Leu Thr Glu Glu Asp Ala Leu Glu Ile
195 200 205

Leu Gln Lys Gly Gln Ile Gly Lys Lys Glu Arg Glu Lys Gln Met Leu
210 215 220

Ala Gln Gly Tyr Pro Ala Tyr Thr Thr Ser Cys Ala Trp Leu Gly Tyr
225 230 235 240

Ser Asp Asp Thr Leu Lys Gln Leu Cys Ala Gln Ala Leu Lys Asp Gly
245 250 255

Trp Thr Arg Phe Lys Val Lys Val Gly Ala Asp Leu Gln Asp Asp Met
260 265 270

Arg Arg Cys Gln Ile Ile Arg Asp Met Ile Gly Pro Glu Lys Thr Leu
275 280 285

Met Met Asp Ala Asn Gln Arg Trp Asp Val Pro Glu Ala Val Glu Trp
290 295 300

241

Met Ser Lys Leu Ala Lys Phe Lys Pro Leu Trp Ile Glu Glu Pro Thr
 305 310 315 320

Ser Pro Asp Asp Ile Leu Gly His Ala Thr Ile Ser Lys Ala Leu Val
 325 330 335

Pro Leu Gly Ile Gly Ile Ala Thr Gly Glu Gln Cys His Asn Arg Val
 340 345 350

Ile Phe Lys Gln Leu Leu Gln Ala Lys Ala Leu Gln Phe Leu Gln Ile
 355 360 365

Asp Ser Cys Arg Leu Gly Ser Val Asn Glu Asn Leu Ser Val Leu Leu
 370 375 380

Met Ala Lys Lys Phe Glu Ile Pro Val Cys Pro His Ala Gly Gly Val
 385 390 395 400

Gly Leu Cys Glu Leu Val Gln His Leu Ile Ile Phe Asp Tyr Ile Ser
 405 410 415

Val Ser Ala Ser Leu Glu Asn Arg Val Cys Glu Tyr Val Asp His Leu
 420 425 430

His Glu His Phe Lys Tyr Pro Val Met Ile Gln Arg Ala Ser Tyr Met
 435 440 445

Pro Pro Lys Asp Pro Gly Tyr Ser Thr Glu Met Lys Glu Glu Ser Val
 450 455 460

Lys Lys His Gln Tyr Pro Asp Gly Glu Val Trp Lys Lys Leu Leu Pro
 465 470 475 480

Ala Gln Glu Asn

<210> 183
 <211> 249
 <212> PRT
 <213> Homo sapien

<400> 183

Arg Met Ala Gly Pro Gly Glu Cys Asp Asp Gly Pro Asp Phe Pro Ser
 1 5 10 15

Trp Arg Gln Glu Arg Leu Arg Gln Phe Lys Val Lys Val Gly Ala Asp
 20 25 30

242

Leu Gln Asp Asp Met Arg Arg Cys Gln Ile Ile Arg Asp Met Ile Gly
 35 40 45

Pro Glu Lys Thr Leu Met Met Asp Ala Asn Gln Arg Trp Asp Val Pro
 50 55 60

Glu Ala Val Glu Trp Met Ser Lys Leu Ala Lys Phe Lys Pro Leu Trp
 65 70 75 80

Ile Glu Glu Pro Thr Ser Pro Asp Asp Ile Leu Gly His Ala Thr Ile
 85 90 95

Ser Lys Ala Leu Val Pro Leu Gly Ile Gly Ile Ala Thr Gly Glu Gln
 100 105 110

Cys His Asn Arg Val Ile Phe Lys Gln Leu Leu Gln Ala Lys Ala Leu
 115 120 125

Gln Phe Leu Gln Ile Asp Ser Cys Arg Leu Gly Ser Val Asn Glu Asn
 130 135 140

Leu Ser Val Leu Leu Met Ala Lys Lys Phe Glu Ile Pro Val Cys Pro
 145 150 155 160

His Ala Gly Gly Val Gly Leu Cys Glu Leu Val Gln His Leu Ile Ile
 165 170 175

Phe Asp Tyr Ile Ser Val Ser Ala Ser Leu Glu Asn Arg Val Cys Glu
 180 185 190

Tyr Val Asp His Leu His Glu His Phe Lys Tyr Pro Val Met Ile Gln
 195 200 205

Arg Ala Ser Tyr Met Pro Pro Lys Asp Pro Gly Tyr Ser Thr Glu Met
 210 215 220

Lys Glu Glu Ser Val Lys Lys His Gln Tyr Pro Asp Gly Glu Val Trp
 225 230 235 240

Lys Lys Leu Leu Pro Ala Gln Glu Asn
 245

<210> 184
 <211> 221
 <212> PRT

243

<213> Homo sapien

<400> 184

Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Trp Ser
1 5 10 15

Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met
20 25 30

Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp
35 40 45

Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser
50 55 60

Ala Asp Ala Met His Thr Asp Pro Asp Tyr Ser Ala Ala Tyr Val Val
65 70 75 80

Ile Glu Thr Asp Ala Glu Asp Gly Ile Lys Gly Cys Gly Ile Thr Phe
85 90 95

Thr Leu Gly Lys Gly Thr Glu Val Val Val Cys Ala Val Asn Ala Leu
100 105 110

Ala His His Val Leu Asn Lys Asp Leu Lys Asp Ile Val Gly Asp Phe
115 120 125

Arg Gly Phe Tyr Arg Gln Leu Thr Ser Asp Gly Gln Leu Arg Trp Ile
130 135 140

Gly Pro Glu Lys Gly Val Val His Leu Ala Thr Ala Ala Val Leu Asn
145 150 155 160

Ala Val Trp Asp Leu Trp Ala Lys Gln Glu Gly Lys Pro Val Trp Lys
165 170 175

Leu Leu Val Asp Met Asp Pro Arg Met Leu Val Ser Cys Ile Asp Phe
180 185 190

Arg Tyr Ile Thr Asp Val Leu Thr Glu Glu Asp Ala Leu Glu Ile Leu
195 200 205

Gln Lys Gly Gln Ile Gly Lys Lys Glu Arg Gly Gly Leu
210 215 220

<210> 185

244

<211> 416
 <212> PRT
 <213> Homo sapien

<400> 185

His Gly Lys Arg Gly Arg His Gly Lys Arg Gly Arg His Gly Met Val
 1 5 10 15

Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala
 20 25 30

Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala
 35 40 45

Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val
 50 55 60

Ser Ala Asp Ala Met His Thr Asp Pro Asp Tyr Ser Ala Ala Tyr Val
 65 70 75 80

Val Ile Glu Thr Asp Ala Glu Asp Gly Ile Lys Gly Cys Gly Ile Thr
 85 90 95

Phe Thr Leu Gly Lys Gly Thr Glu Val Val Val Cys Ala Val Asn Ala
 100 105 110

Leu Ala His His Val Leu Asn Lys Asp Leu Lys Asp Ile Val Gly Asp
 115 120 125

Phe Arg Gly Phe Tyr Arg Gln Leu Thr Ser Asp Gly Gln Leu Arg Trp
 130 135 140

Ile Gly Pro Glu Lys Gly Val Val His Leu Ala Thr Ala Ala Val Leu
 145 150 155 160

Asn Ala Val Trp Asp Leu Trp Ala Lys Gln Glu Gly Lys Pro Val Trp
 165 170 175

Lys Leu Leu Val Asp Met Asp Pro Arg Met Leu Val Ser Cys Ile Asp
 180 185 190

Phe Arg Tyr Ile Thr Asp Val Leu Thr Glu Glu Asp Ala Leu Glu Ile
 195 200 205

Leu Gln Lys Gly Gln Ile Gly Lys Lys Glu Arg Glu Lys Gln Met Leu
 210 215 220

245

Ala Gln Gly Tyr Pro Ala Tyr Thr Thr Ser Cys Ala Trp Leu Gly Tyr
 225 230 235 240

Ser Asp Asp Thr Leu Lys Gln Leu Cys Ala Gln Ala Leu Lys Asp Gly
 245 250 255

Trp Thr Arg Phe Lys Val Lys Val Gly Ala Asp Leu Gln Asp Asp Met
 260 265 270

Arg Arg Cys Gln Ile Ile Arg Asp Met Ile Gly Pro Glu Lys Thr Leu
 275 280 285

Met Met Asp Ala Asn Gln Arg Trp Asp Val Pro Glu Ala Val Glu Trp
 290 295 300

Met Ser Lys Leu Ala Lys Phe Lys Pro Leu Trp Ile Glu Glu Pro Thr
 305 310 315 320

Ser Pro Asp Asp Ile Leu Gly His Ala Thr Ile Ser Lys Ala Leu Val
 325 330 335

Pro Leu Gly Ile Gly Ile Ala Thr Gly Glu Gln Val Ser Asp Ala Pro
 340 345 350

Asn Arg Trp Met Thr Ser Pro Trp Gly Gln Tyr Thr Leu Thr Ser Asp
 355 360 365

Arg Gly His Ser Cys Val Leu Gly Ser Ile Thr Cys Cys Thr Leu Ser
 370 375 380

Trp Glu Ile Phe Ile Ile Leu Glu Thr Gly Ser Phe Tyr Gln Ser Leu
 385 390 395 400

Glu Ser Asp Ile Glu Lys Val Cys Gly Tyr Phe Ser Asn Leu Tyr Asp
 405 410 415

<210> 186
 <211> 415
 <212> PRT
 <213> Homo sapien

<400> 186

Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Trp Ser
 1 5 10 15

Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met

246

20	25	30
Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp		
35	40	45
Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser		
50	55	60
Ala Asp Ala Met His Thr Asp Pro Asp Tyr Ser Ala Ala Tyr Val Val		
65	70	75
Ile Glu Thr Asp Ala Glu Asp Gly Ile Lys Gly Cys Gly Ile Thr Phe		
85	90	95
Thr Leu Gly Lys Gly Thr Glu Val Val Val Cys Ala Val Asn Ala Leu		
100	105	110
Ala His His Val Leu Asn Lys Asp Leu Lys Asp Ile Val Gly Asp Phe		
115	120	125
Arg Gly Phe Tyr Arg Gln Leu Thr Ser Asp Gly Gln Leu Arg Trp Ile		
130	135	140
Gly Pro Glu Lys Gly Val Val His Leu Ala Thr Ala Ala Val Leu Asn		
145	150	155
Ala Val Trp Asp Leu Trp Ala Lys Gln Glu Gly Lys Pro Val Trp Lys		
165	170	175
Leu Leu Val Asp Met Asp Pro Arg Met Leu Val Ser Cys Ile Asp Phe		
180	185	190
Arg Tyr Ile Thr Asp Val Leu Thr Glu Glu Asp Ala Leu Glu Ile Leu		
195	200	205
Gln Lys Gly Gln Ile Gly Lys Lys Glu Arg Glu Lys Gln Met Leu Ala		
210	215	220
Gln Gly Tyr Pro Ala Tyr Thr Thr Ser Cys Ala Trp Leu Gly Tyr Ser		
225	230	235
Asp Asp Thr Leu Lys Gln Leu Cys Ala Gln Ala Leu Lys Asp Gly Trp		
245	250	255
Thr Arg Phe Lys Val Lys Val Gly Ala Asp Leu Gln Asp Asp Met Arg		
260	265	270

247

Arg Cys Gln Ile Ile Arg Asp Met Ile Gly Pro Glu Lys Thr Leu Met
 275 280 285

Met Asp Ala Asn Gln Arg Trp Asp Val Pro Glu Ala Val Glu Trp Met
 290 295 300

Ser Lys Leu Ala Lys Phe Lys Pro Leu Trp Ile Glu Glu Pro Thr Ser
 305 310 315 320

Pro Asp Asp Ile Leu Gly His Ala Thr Ile Ser Lys Ala Leu Val Pro
 325 330 335

Leu Gly Ile Gly Ile Ala Thr Gly Glu Gln Val Ser Asp Ala Pro Asn
 340 345 350

Arg Trp Met Thr Ser Pro Trp Gly Gln Tyr Thr Leu Thr Ser Asp Arg
 355 360 365

Gly His Ser Cys Val Leu Gly Ser Ile Thr Cys Cys Thr Leu Ser Trp
 370 375 380

Glu Ile Phe Ile Ile Leu Glu Thr Gly Ser Phe Tyr Gln Ser Leu Glu
 385 390 395 400

Ser Asp Ile Glu Lys Val Cys Gly Tyr Phe Ser Asn Leu Tyr Asp
 405 410 415

<210> 187
 <211> 484
 <212> PRT
 <213> Homo sapien

<400> 187

His Gly Lys Arg Gly Arg His Gly Lys Arg Gly Arg His Gly Met Val
 1 5 10 15

Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala
 20 25 30

Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala
 35 40 45

Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val
 50 55 60

248

Ser Ala Asp Ala Met His Thr Asp Pro Asp Tyr Ser Ala Ala Tyr Val
65 70 75 80

Val Ile Glu Thr Asp Ala Glu Asp Gly Ile Lys Gly Cys Gly Ile Thr
85 90 95

Phe Thr Leu Gly Lys Gly Thr Glu Val Val Val Cys Ala Val Asn Ala
100 105 110

Leu Ala His His Val Leu Asn Lys Asp Leu Lys Asp Ile Val Gly Asp
115 120 125

Phe Arg Gly Phe Tyr Arg Gln Leu Thr Ser Asp Gly Gln Leu Arg Trp
130 135 140

Ile Gly Pro Glu Lys Gly Val Val His Leu Ala Thr Ala Ala Val Leu
145 150 155 160

Asn Ala Val Trp Asp Leu Trp Ala Lys Gln Glu Gly Lys Pro Val Trp
165 170 175

Lys Leu Leu Val Asp Met Asp Pro Arg Met Leu Val Ser Cys Ile Asp
180 185 190

Phe Arg Tyr Ile Thr Asp Val Leu Thr Glu Glu Asp Ala Leu Glu Ile
195 200 205

Leu Gln Lys Gly Gln Ile Gly Lys Lys Glu Arg Glu Lys Gln Met Leu
210 215 220

Ala Gln Gly Tyr Pro Ala Tyr Thr Thr Ser Cys Ala Trp Leu Gly Tyr
225 230 235 240

Ser Asp Asp Thr Leu Lys Gln Leu Cys Ala Gln Ala Leu Lys Asp Gly
245 250 255

Trp Thr Arg Phe Lys Val Lys Val Gly Ala Asp Leu Gln Asp Asp Met
260 265 270

Arg Arg Cys Gln Ile Ile Arg Asp Met Ile Gly Pro Glu Lys Thr Leu
275 280 285

Met Met Asp Ala Asn Gln Arg Trp Asp Val Pro Glu Ala Val Glu Trp
290 295 300

Met Ser Lys Leu Ala Lys Phe Lys Pro Leu Trp Ile Glu Glu Pro Thr

Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala
20 25 30

250

Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala
 35 40 45

Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val
 50 55 60

Ser Ala Asp Ala Met His Thr Asp Pro Asp Tyr Ser Ala Ala Tyr Val
 65 70 75 80

Val Ile Glu Thr Asp Ala Glu Asp Gly Ile Lys Gly Cys Gly Ile Thr
 85 90 95

Phe Thr Leu Gly Lys Gly Thr Glu Val Val Val Cys Ala Val Asn Ala
 100 105 110

Leu Ala His His Val Leu Asn Lys Asp Leu Lys Asp Ile Val Gly Asp
 115 120 125

Phe Arg Gly Phe Tyr Arg Gln Leu Thr Ser Asp Gly Gln Leu Arg Trp
 130 135 140

Ile Gly Pro Glu Lys Gly Val Val His Leu Ala Thr Ala Ala Val Leu
 145 150 155 160

Asn Ala Val Trp Asp Leu Trp Ala Lys Gln Glu Gly Lys Pro Val Trp
 165 170 175

Lys Leu Leu Val Asp Met Asp Pro Arg Met Leu Val Ser Cys Ile Asp
 180 185 190

Phe Arg Tyr Ile Thr Asp Val Leu Thr Glu Glu Asp Ala Leu Glu Ile
 195 200 205

Leu Gln Lys Gly Gln Ile Gly Lys Lys Glu Arg Glu Lys Gln Met Leu
 210 215 220

Ala Gln Gly Tyr Pro Ala Tyr Thr Thr Ser Cys Ala Trp Leu Gly Tyr
 225 230 235 240

Ser Asp Asp Thr Leu Lys Gln Leu Cys Ala Gln Ala Leu Lys Asp Gly
 245 250 255

Trp Thr Arg Phe Lys Val Lys Val Gly Ala Asp Leu Gln Asp Asp Met
 260 265 270

251

Arg Arg Cys Gln Ile Ile Arg Asp Met Ile Gly Pro Glu Lys Thr Leu
 275 280 285

Met Met Asp Ala Asn Gln Arg Trp Asp Val Pro Glu Ala Val Glu Trp
 290 295 300

Met Ser Lys Leu Ala Lys Phe Lys Pro Leu Trp Ile Glu Glu Pro Thr
 305 310 315 320

Ser Pro Asp Asp Ile Leu Gly His Ala Thr Ile Ser Lys Ala Leu Val
 325 330 335

Pro Leu Gly Ile Gly Ile Ala Thr Gly Glu Gln Gly Val
 340 345

<210> 189
 <211> 305
 <212> PRT
 <213> Homo sapien

<400> 189

Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala
 1 5 10 15

Asp Ala Met Val Ser Ala Asp Ala Met His Thr Asp Pro Asp Tyr Ser
 20 25 30

Ala Ala Tyr Val Val Ile Glu Thr Asp Ala Glu Asp Gly Ile Lys Gly
 35 40 45

Cys Gly Ile Thr Phe Thr Leu Gly Lys Gly Thr Glu Val Val Val Cys
 50 55 60

Ala Val Asn Ala Leu Ala His His Val Leu Asn Lys Asp Leu Lys Asp
 65 70 75 80

Ile Val Gly Asp Phe Arg Gly Phe Tyr Arg Gln Leu Thr Ser Asp Gly
 85 90 95

Gln Leu Arg Trp Ile Gly Pro Glu Lys Gly Val Val His Leu Ala Thr
 100 105 110

Ala Ala Val Leu Asn Ala Val Trp Asp Leu Trp Ala Lys Gln Glu Gly
 115 120 125

Lys Pro Val Trp Lys Leu Leu Val Asp Met Asp Pro Arg Met Leu Val
 130 135 140

252

Ser Cys Ile Asp Phe Arg Tyr Ile Thr Asp Val Leu Thr Glu Glu Asp
 145 150 155 160

Ala Leu Glu Ile Leu Gln Lys Gly Gln Ile Gly Lys Lys Glu Arg Glu
 165 170 175

Lys Gln Met Leu Ala Gln Gly Tyr Pro Ala Tyr Thr Thr Ser Cys Ala
 180 185 190

Trp Leu Gly Tyr Ser Asp Asp Thr Leu Lys Gln Leu Cys Ala Gln Ala
 195 200 205

Leu Lys Asp Gly Trp Thr Arg Phe Lys Val Lys Val Gly Ala Asp Leu
 210 215 220

Gln Asp Asp Met Arg Arg Cys Gln Ile Ile Arg Asp Met Ile Gly Pro
 225 230 235 240

Glu Lys Thr Leu Met Met Asp Ala Asn Gln Arg Trp Asp Val Pro Glu
 245 250 255

Ala Val Glu Trp Met Ser Lys Leu Ala Lys Phe Lys Pro Leu Trp Ile
 260 265 270

Glu Glu Pro Thr Ser Pro Asp Asp Ile Leu Gly His Ala Thr Ile Ser
 275 280 285

Lys Ala Leu Val Pro Leu Gly Ile Gly Ile Ala Thr Gly Glu Gln Gly
 290 295 300

Val
 305

<210> 190
 <211> 484
 <212> PRT
 <213> Homo sapien

<400> 190

His Gly Lys Arg Gly Arg His Gly Lys Arg Gly Arg His Gly Met Val
 1 5 10 15

Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala
 20 25 30

253

Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala
 35 40 45

Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val
 50 55 60

Ser Ala Asp Ala Met His Thr Asp Pro Asp Tyr Ser Ala Ala Tyr Val
 65 70 75 80

Val Ile Glu Thr Asp Ala Glu Asp Gly Ile Lys Gly Cys Gly Ile Thr
 85 90 95

Phe Thr Leu Gly Lys Gly Thr Glu Val Val Val Cys Ala Val Asn Ala
 100 105 110

Leu Ala His His Val Leu Asn Lys Asp Leu Lys Asp Ile Val Gly Asp
 115 120 125

Phe Arg Gly Phe Tyr Arg Gln Leu Thr Ser Asp Gly Gln Leu Arg Trp
 130 135 140

Ile Gly Pro Glu Lys Gly Val Val His Leu Ala Thr Ala Ala Val Leu
 145 150 155 160

Asn Ala Val Trp Asp Leu Trp Ala Lys Gln Glu Gly Lys Pro Val Trp
 165 170 175

Lys Leu Leu Val Asp Met Asp Pro Arg Met Leu Val Ser Cys Ile Asp
 180 185 190

Phe Arg Tyr Ile Thr Asp Val Leu Thr Glu Glu Asp Ala Leu Glu Ile
 195 200 205

Leu Gln Lys Gly Gln Ile Gly Lys Lys Glu Arg Glu Lys Gln Met Leu
 210 215 220

Ala Gln Gly Tyr Pro Ala Tyr Thr Thr Ser Cys Ala Trp Leu Gly Tyr
 225 230 235 240

Ser Asp Asp Thr Leu Lys Gln Leu Cys Ala Gln Ala Leu Lys Asp Gly
 245 250 255

Trp Thr Arg Phe Lys Val Lys Val Gly Ala Asp Leu Gln Asp Asp Met
 260 265 270

Arg Arg Cys Gln Ile Ile Arg Asp Met Ile Gly Pro Glu Lys Thr Leu

254

275

280

285

Met Met Asp Ala Asn Gln Arg Trp Asp Val Pro Glu Ala Val Glu Trp
 290 295 300

Met Ser Lys Leu Ala Lys Phe Lys Pro Leu Trp Ile Glu Glu Pro Thr
 305 310 315 320

Ser Pro Asp Asp Ile Leu Gly His Ala Thr Ile Ser Lys Ala Leu Val
 325 330 335

Pro Leu Gly Ile Gly Ile Ala Thr Gly Glu Gln Cys His Asn Arg Val
 340 345 350

Ile Phe Lys Gln Leu Leu Gln Ala Lys Ala Leu Gln Phe Leu Gln Ile
 355 360 365

Asp Ser Cys Arg Leu Gly Ser Val Asn Glu Asn Leu Ser Val Leu Leu
 370 375 380

Met Ala Lys Lys Phe Glu Ile Pro Val Cys Pro His Ala Gly Gly Val
 385 390 395 400

Gly Leu Cys Glu Leu Val Gln His Leu Ile Ile Phe Asp Tyr Ile Ser
 405 410 415

Val Ser Ala Ser Leu Glu Asn Arg Val Cys Glu Tyr Val Asp His Leu
 420 425 430

His Glu His Phe Lys Tyr Pro Val Met Ile Gln Arg Ala Ser Tyr Met
 435 440 445

Pro Pro Lys Asp Pro Gly Tyr Ser Thr Glu Met Lys Glu Glu Ser Val
 450 455 460

Lys Lys His Gln Tyr Pro Asp Gly Glu Val Trp Lys Lys Leu Leu Pro
 465 470 475 480

Ala Gln Glu Asn

<210> 191

<211> 484

<212> PRT

<213> Homo sapien

<400> 191

255

His Gly Lys Arg Gly Arg His Gly Lys Arg Gly Arg His Gly Met Val
 1 5 10 15

Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala
 20 25 30

Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala
 35 40 45

Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val
 50 55 60

Ser Ala Asp Ala Met His Thr Asp Pro Asp Tyr Ser Ala Ala Tyr Val
 65 70 75 80

Val Ile Glu Thr Asp Ala Glu Asp Gly Ile Lys Gly Cys Gly Ile Thr
 85 90 95

Phe Thr Leu Gly Lys Gly Thr Glu Val Val Val Cys Ala Val Asn Ala
 100 105 110

Leu Ala His His Val Leu Asn Lys Asp Leu Lys Asp Ile Val Gly Asp
 115 120 125

Phe Arg Gly Phe Tyr Arg Gln Leu Thr Ser Asp Gly Gln Leu Arg Trp
 130 135 140

Ile Gly Pro Glu Lys Gly Val Val His Leu Ala Thr Ala Ala Val Leu
 145 150 155 160

Asn Ala Val Trp Asp Leu Trp Ala Lys Gln Glu Gly Lys Pro Val Trp
 165 170 175

Lys Leu Leu Val Asp Met Asp Pro Arg Met Leu Val Ser Cys Ile Asp
 180 185 190

Phe Arg Tyr Ile Thr Asp Val Leu Thr Glu Glu Asp Ala Leu Glu Ile
 195 200 205

Leu Gln Lys Gly Gln Ile Gly Lys Lys Glu Arg Glu Lys Gln Met Leu
 210 215 220

Ala Gln Gly Tyr Pro Ala Tyr Thr Thr Ser Cys Ala Trp Leu Gly Tyr
 225 230 235 240

256

Ser Asp Asp Thr Leu Lys Gln Leu Cys Ala Gln Ala Leu Lys Asp Gly
 245 250 255

Trp Thr Arg Phe Lys Val Lys Val Gly Ala Asp Leu Gln Asp Asp Met
 260 265 270

Arg Arg Cys Gln Ile Ile Arg Asp Met Ile Gly Pro Glu Lys Thr Leu
 275 280 285

Met Met Asp Ala Asn Gln Arg Trp Asp Val Pro Glu Ala Val Glu Trp
 290 295 300

Met Ser Lys Leu Ala Lys Phe Lys Pro Leu Trp Ile Glu Glu Pro Thr
 305 310 315 320

Ser Pro Asp Asp Ile Leu Gly His Ala Thr Ile Ser Lys Ala Leu Val
 325 330 335

Pro Leu Gly Ile Gly Ile Ala Thr Gly Glu Gln Cys His Asn Arg Val
 340 345 350

Ile Phe Lys Gln Leu Leu Gln Ala Lys Ala Leu Gln Phe Leu Gln Ile
 355 360 365

Asp Ser Cys Arg Leu Gly Ser Val Asn Glu Asn Leu Ser Val Leu Leu
 370 375 380

Met Ala Lys Lys Phe Glu Ile Pro Val Cys Pro His Ala Gly Gly Val
 385 390 395 400

Gly Leu Cys Glu Leu Val Gln His Leu Ile Ile Phe Asp Tyr Ile Ser
 405 410 415

Val Ser Ala Ser Leu Glu Asn Arg Val Cys Glu Tyr Val Asp His Leu
 420 425 430

His Glu His Phe Lys Tyr Pro Val Met Ile Gln Arg Ala Ser Tyr Met
 435 440 445

Pro Pro Lys Asp Pro Gly Tyr Ser Thr Glu Met Lys Glu Glu Ser Val
 450 455 460

Lys Lys His Gln Tyr Pro Asp Gly Glu Val Trp Lys Lys Leu Leu Pro
 465 470 475 480

Ala Gln Glu Asn

257

<210> 192
 <211> 484
 <212> PRT
 <213> Homo sapien

<400> 192

His Gly Lys Arg Gly Arg His Gly Lys Arg Gly Arg His Gly Met Val
 1 5 10 15

Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala
 20 25 30

Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala
 35 40 45

Asp Ala Met Val Ser Ala Asp Ala Met Val Ser Ala Asp Ala Met Val
 50 55 60

Ser Ala Asp Ala Met His Thr Asp Pro Asp Tyr Ser Ala Ala Tyr Val
 65 70 75 80

Val Ile Glu Thr Asp Ala Glu Asp Gly Ile Lys Gly Cys Gly Ile Thr
 85 90 95

Phe Thr Leu Gly Lys Gly Thr Glu Val Val Val Cys Ala Val Asn Ala
 100 105 110

Leu Ala His His Val Leu Asn Lys Asp Leu Lys Asp Ile Val Gly Asp
 115 120 125

Phe Arg Gly Phe Tyr Arg Gln Leu Thr Ser Asp Gly Gln Leu Arg Trp
 130 135 140

Ile Gly Pro Glu Lys Gly Val Val His Leu Ala Thr Ala Ala Val Leu
 145 150 155 160

Asn Ala Val Trp Asp Leu Trp Ala Lys Gln Glu Gly Lys Pro Val Trp
 165 170 175

Lys Leu Leu Val Asp Met Asp Pro Arg Met Leu Val Ser Cys Ile Asp
 180 185 190

Phe Arg Tyr Ile Thr Asp Val Leu Thr Glu Glu Asp Ala Leu Glu Ile
 195 200 205

258

Leu Gln Lys Gly Gln Ile Gly Lys Lys Glu Arg Glu Lys Gln Met Leu
 210 215 220

Ala Gln Gly Tyr Pro Ala Tyr Thr Thr Ser Cys Ala Trp Leu Gly Tyr
 225 230 235 240

Ser Asp Asp Thr Leu Lys Gln Leu Cys Ala Gln Ala Leu Lys Asp Gly
 245 250 255

Trp Thr Arg Phe Lys Val Lys Val Gly Ala Asp Leu Gln Asp Asp Met
 260 265 270

Arg Arg Cys Gln Ile Ile Arg Asp Met Ile Gly Pro Glu Lys Thr Leu
 275 280 285

Met Met Asp Ala Asn Gln Arg Trp Asp Val Pro Glu Ala Val Glu Trp
 290 295 300

Met Ser Lys Leu Ala Lys Phe Lys Pro Leu Trp Ile Glu Glu Pro Thr
 305 310 315 320

Ser Pro Asp Asp Ile Leu Gly His Ala Thr Ile Ser Lys Ala Leu Val
 325 330 335

Pro Leu Gly Ile Gly Ile Ala Thr Gly Glu Gln Cys His Asn Arg Val
 340 345 350

Ile Phe Lys Gln Leu Leu Gln Ala Lys Ala Leu Gln Phe Leu Gln Ile
 355 360 365

Asp Ser Cys Arg Leu Gly Ser Val Asn Glu Asn Leu Ser Val Leu Leu
 370 375 380

Met Ala Lys Lys Phe Glu Ile Pro Val Cys Pro His Ala Gly Gly Val
 385 390 395 400

Gly Leu Cys Glu Leu Val Gln His Leu Ile Ile Phe Asp Tyr Ile Ser
 405 410 415

Val Ser Ala Ser Leu Glu Asn Arg Val Cys Glu Tyr Val Asp His Leu
 420 425 430

His Glu His Phe Lys Tyr Pro Val Met Ile Gln Arg Ala Ser Tyr Met
 435 440 445

259

Pro Pro Lys Asp Pro Gly Tyr Ser Thr Glu Met Lys Glu Glu Ser Val
 450 455 460

Lys Lys His Gln Tyr Pro Asp Gly Glu Val Trp Lys Lys Leu Leu Pro
 465 470 475 480

Ala Gln Glu Asn

<210> 193
 <211> 138
 <212> PRT
 <213> Homo sapien

<400> 193

Trp Ile Val Val Ala Ala Arg Tyr Arg Ile Arg Leu Gly Leu Tyr Leu
 1 5 10 15

Thr Leu Ala Ser Glu Val Tyr Tyr Thr Arg Leu Gly Asn Asp Phe His
 20 25 30

Thr Asn Lys Arg Val Cys Glu Glu Ile Ala Ile Ile Pro Ser Lys Lys
 35 40 45

Leu Arg Asn Lys Ile Ala Gly Tyr Val Thr His Leu Met Lys Arg Ile
 50 55 60

Gln Arg Gly Pro Val Arg Gly Ile Ser Ile Lys Leu Gln Glu Glu Glu
 65 70 75 80

Arg Glu Arg Arg Asp Asn Tyr Val Pro Glu Val Ser Ala Leu Asp Gln
 85 90 95

Glu Ile Ile Glu Val Asp Pro Asp Thr Lys Glu Met Leu Lys Leu Leu
 100 105 110

Asp Phe Gly Ser Leu Ser Asn Leu Gln Val Thr Gln Pro Thr Val Gly
 115 120 125

Met Asn Phe Lys Thr Pro Arg Gly Pro Val
 130 135

<210> 194
 <211> 386
 <212> PRT
 <213> Homo sapien

<400> 194

260

Met Pro Trp Ala Met Ile Trp Asp Phe Thr Glu Pro Val Cys Arg Gly
 1 5 10 15

Cys Val Asn Tyr Glu Gly Ala Asp Arg Val Glu Phe Val Ile Glu Thr
 20 25 30

Ala Arg Gln Leu Lys Arg Ala His Gly Cys Phe Pro Glu Gly Arg Ser
 35 40 45

Pro Pro Gly Ala Ala Ala Ser Ala Ala Ala Lys Pro Pro Pro Leu Ser
 50 55 60

Ala Lys Asp Ile Leu Leu Gln Gln Gln Gln Leu Gly His Gly Gly
 65 70 75 80

Pro Glu Ala Ala Pro Arg Ala Pro Gln Ala Leu Glu Arg Tyr Pro Leu
 85 90 95

Ala Ala Ala Ala Glu Arg Pro Pro Arg Leu Gly Ser Asp Phe Gly Ser
 100 105 110

Ser Arg Pro Ala Ala Ser Leu Ala Gln Pro Pro Thr Pro Gln Pro Pro
 115 120 125

Pro Val Asn Gly Ile Leu Val Pro Asn Gly Phe Ser Lys Leu Glu Glu
 130 135 140

Pro Pro Glu Leu Asn Arg Gln Ser Pro Asn Pro Arg Arg Gly His Ala
 145 150 155 160

Val Pro Pro Thr Leu Val Pro Leu Met Asn Gly Ser Ala Thr Pro Leu
 165 170 175

Pro Thr Ala Leu Gly Leu Gly Gly Arg Ala Ala Ala Ser Leu Ala Ala
 180 185 190

Val Ser Gly Thr Ala Ala Ala Ser Leu Gly Ser Ala Gln Pro Thr Asp
 195 200 205

Leu Gly Ala His Lys Arg Pro Ala Ser Val Ser Ser Ser Ala Ala Val
 210 215 220

Glu His Glu Gln Arg Glu Ala Ala Ala Lys Glu Lys Gln Pro Pro Pro
 225 230 235 240

261

Pro Ala His Arg Gly Pro Ala Asp Ser Leu Ser Thr Ala Ala Gly Ala
245 250 255

Ala Glu Leu Ser Ala Glu Gly Ala Gly Lys Ser Arg Gly Ser Gly Glu
260 265 270

Gln Asp Trp Val Asn Arg Pro Lys Thr Val Arg Asp Thr Leu Leu Ala
275 280 285

Leu His Gln His Gly His Ser Gly Pro Phe Glu Ser Lys Phe Lys Lys
290 295 300

Glu Pro Ala Leu Thr Ala Gly Arg Leu Leu Gly Phe Glu Ala Asn Gly
305 310 315 320

Ala Asn Gly Ser Lys Ala Gly Arg Gly Gly Cys Glu Val Arg Gly Ser
325 330 335

Arg Gly Glu Lys Gly Thr Glu Ser Arg Gly Arg Val Val Leu Trp Ile
340 345 350

His His Phe Thr Pro Ala Gln Lys Gln Gln Thr Pro His Phe Leu Ile
355 360 365

Cys Leu Arg Arg Asn Gln Cys Leu Val Ala Thr Cys Ser Cys Ala Glu
370 375 380

Ala Ala
385

```
<210> 195
<211> 492
<212> PRT
<213> Homo sapien
```

<400> 195

Met Pro Trp Ala Met Ile Trp Asp Phe Thr Glu Pro Val Cys Arg Gly
1 5 10 15

Cys Val Asn Tyr Glu Gly Ala Asp Arg Val Glu Phe Val Ile Glu Thr
20 25 30

Ala Arg Gln Leu Lys Arg Ala His Gly Cys Phe Pro Glu Gly Arg Ser
35 40 45

Pro Pro Gly Ala Ala Ala Ser Ala Ala Ala Lys Pro Pro Pro Leu Ser
50 55 60

262

Ala Lys Asp Ile Leu Leu Gln Gln Gln Gln Gln Leu Gly His Gly Gly
 65 70 75 80

Pro Glu Ala Ala Pro Arg Ala Pro Gln Ala Leu Glu Arg Tyr Pro Leu
 85 90 95

Ala Ala Ala Ala Glu Arg Pro Pro Arg Leu Gly Ser Asp Phe Gly Ser
 100 105 110

Ser Arg Pro Ala Ala Ser Leu Ala Gln Pro Pro Thr Pro Gln Pro Pro
 115 120 125

Pro Val Asn Gly Ile Leu Val Pro Asn Gly Phe Ser Lys Leu Glu Glu
 130 135 140

Pro Pro Glu Leu Asn Arg Gln Ser Pro Asn Pro Arg Arg Gly His Ala
 145 150 155 160

Val Pro Pro Thr Leu Val Pro Leu Met Asn Gly Ser Ala Thr Pro Leu
 165 170 175

Pro Thr Ala Leu Gly Leu Gly Gly Arg Ala Ala Ala Ser Leu Ala Ala
 180 185 190

Val Ser Gly Thr Ala Ala Ala Ser Leu Gly Ser Ala Gln Pro Thr Asp
 195 200 205

Leu Gly Ala His Lys Arg Pro Ala Ser Val Ser Ser Ser Ala Ala Val
 210 215 220

Glu His Glu Gln Arg Glu Ala Ala Ala Lys Glu Lys Gln Pro Pro Pro
 225 230 235 240

Pro Ala His Arg Gly Pro Ala Asp Ser Leu Ser Thr Ala Ala Gly Ala
 245 250 255

Ala Glu Leu Ser Ala Glu Gly Ala Gly Lys Ser Arg Gly Ser Gly Glu
 260 265 270

Gln Asp Trp Val Asn Arg Pro Lys Thr Val Arg Asp Thr Leu Leu Ala
 275 280 285

Leu His Gln His Gly His Ser Gly Pro Phe Glu Ser Lys Phe Lys Lys
 290 295 300

263

Glu Pro Ala Leu Thr Ala Val Ala Arg Thr Ala Arg Lys Arg Lys Pro
 305 310 315 320

Ser Pro Glu Pro Glu Gly Glu Val Gly Pro Pro Lys Ile Asn Gly Glu
 325 330 335

Ala Gln Pro Trp Leu Ser Thr Ser Thr Glu Gly Leu Lys Ile Pro Met
 340 345 350

Thr Pro Thr Ser Ser Phe Val Ser Pro Pro Pro Pro Thr Ala Ser Pro
 355 360 365

His Ser Asn Arg Thr Thr Pro Pro Glu Ala Ala Gln Asn Gly Gln Ser
 370 375 380

Pro Met Ala Ala Leu Ile Leu Val Ala Asp Asn Ala Gly Gly Ser His
 385 390 395 400

Ala Ser Lys Asp Ala Asn Gln Val His Pro Leu Trp Gln Pro Val Pro
 405 410 415

Arg Cys Ala Ala Pro Ser Ala Thr Ser Gly Trp Arg Thr Pro Ile Leu
 420 425 430

Cys Ser Ala Arg Pro Ser Leu Arg Thr Ser Ser Ala Ser Leu Ala Pro
 435 440 445

Asp Lys Ala Ser Asn Ser Arg Glu Leu Val Glu Arg Ser Ile Val Pro
 450 455 460

Val Gly Lys Asn Ala Leu Leu Trp Ala Pro Met Ser Pro Gly Pro Leu
 465 470 475 480

Cys Lys Gly Lys Leu Gln Pro Ser Leu Leu Glu Met
 485 490

<210> 196
 <211> 358
 <212> PRT
 <213> Homo sapien

<400> 196

Met Ser Gly Val Arg Pro Pro Ile Met Asn Gly Pro Leu His Pro Arg
 1 5 10 15

Pro Leu Val Ala Leu Leu Asp Gly Arg Asp Cys Thr Val Glu Met Pro

264

20

25

30

Ile Leu Lys Asp Val Ala Thr Val Ala Phe Cys Asp Ala Gln Ser Thr
 35 40 45

Gln Glu Ile His Glu Lys Val Leu Asn Glu Ala Val Gly Ala Leu Met
 50 55 60

Tyr His Thr Ile Thr Leu Thr Arg Glu Asp Leu Glu Lys Phe Lys Ala
 65 70 75 80

Leu Arg Ile Ile Val Arg Ile Gly Ser Gly Phe Asp Asn Ile Asp Ile
 85 90 95

Lys Ser Ala Gly Asp Leu Gly Ile Ala Val Cys Asn Val Pro Ala Ala
 100 105 110

Ser Val Glu Glu Thr Ala Asp Ser Thr Leu Cys His Ile Leu Asn Leu
 115 120 125

Tyr Arg Arg Ala Thr Trp Leu His Gln Ala Leu Arg Glu Gly Thr Arg
 130 135 140

Val Gln Ser Val Glu Gln Ile Arg Glu Val Ala Ser Gly Ala Ala Arg
 145 150 155 160

Ile Arg Gly Glu Thr Leu Gly Ile Ile Gly Leu Gly Arg Val Gly Gln
 165 170 175

Ala Val Ala Leu Arg Ala Lys Ala Phe Gly Phe Asn Val Leu Phe Tyr
 180 185 190

Asp Pro Tyr Leu Ser Asp Gly Val Glu Arg Ala Leu Gly Leu Gln Arg
 195 200 205

Val Ser Thr Leu Gln Asp Leu Leu Phe His Ser Asp Cys Val Thr Leu
 210 215 220

His Cys Gly Leu Asn Glu His Asn His His Leu Ile Asn Asp Phe Thr
 225 230 235 240

Val Lys Gln Met Arg Gln Gly Ala Phe Leu Val Asn Thr Ala Arg Gly
 245 250 255

Gly Leu Val Asp Glu Lys Ala Leu Ala Gln Ala Leu Lys Glu Gly Arg
 260 265 270

265

Ile Arg Gly Ala Ala Leu Asp Val His Glu Ser Glu Pro Phe Ser Phe
 275 280 285

Ser Gln Gly Pro Leu Lys Asp Ala Pro Asn Leu Ile Cys Thr Pro His
 290 295 300

Ala Ala Trp Tyr Ser Glu Gln Ala Ser Ile Glu Met Arg Glu Glu Ala
 305 310 315 320

Ala Arg Glu Ile Arg Arg Ala Ile Thr Gly Arg Ile Pro Asp Ser Leu
 325 330 335

Lys Asn Phe Cys Pro Val Ser Phe Ala Phe Leu Val Lys Gln Lys Lys
 340 345 350

Ser Val Val Ile Leu Pro
 355

<210> 197
 <211> 364
 <212> PRT
 <213> Homo sapien

<400> 197

Met Gly Pro Gly His Gly Val Met Ala Ser Arg Pro Asp Leu Gln Pro
 1 5 10 15

Leu Gln His Leu Gly Thr Pro Gly Ser Pro Gly Leu Asp Val Gln Pro
 20 25 30

Gln Glu Glu Thr Pro Pro Gln Gly Gln Tyr Gln Pro Ala Ala Pro Gly
 35 40 45

Ala Thr Asp Pro Leu Ala Gly Arg Gly Gln Ala Ala Cys Pro Pro Ile
 50 55 60

Arg Ala Pro Pro Thr Arg Asp Leu Glu Ile Lys Ser Leu Gly Leu Pro
 65 70 75 80

His Pro Pro Leu Ser Gly Ala Pro Gly Val Ser Asp Gly Pro Gly Ala
 85 90 95

Val Leu Leu Ser Ser Ala Ser Leu Pro Ser Arg Ala Gly Pro Trp Gly
 100 105 110

266

Leu Trp Phe Pro Gly Arg Ala Pro His Arg Gly Phe Gln Cys Gln Pro
 115 120 125

Pro Pro Leu Arg Thr Gln Pro Gln His Ser Gly Cys Thr Asp His Ala
 130 135 140

Cys Ala Val Pro Ser Phe Ser Gln Gly Pro Leu Lys Asp Ala Pro Asn
 145 150 155 160

Leu Ile Cys Thr Pro His Ala Ala Trp Tyr Ser Glu Gln Ala Ser Ile
 165 170 175

Glu Met Arg Glu Glu Ala Ala Arg Glu Ile Arg Arg Ala Ile Thr Gly
 180 185 190

Arg Ile Pro Asp Ser Leu Lys Asn Cys Val Asn Lys Asp His Leu Thr
 195 200 205

Ala Ala Thr His Trp Ala Ser Met Asp Pro Ala Val Val His Pro Glu
 210 215 220

Leu Asn Gly Ala Ala Tyr Ser Arg Gly Thr Leu Arg Ala Trp Trp Ala
 225 230 235 240

Trp Pro Pro Leu Ala Ser Gln Leu Leu Trp Lys Val Ser Ser Pro Ala
 245 250 255

Pro Cys Pro Cys Pro Thr Ala Cys Pro Leu Trp Pro Thr Arg Pro Thr
 260 265 270

Pro Leu Leu Leu Ala Lys Pro Ser Ser Pro Arg Arg Ile Glu Thr Thr
 275 280 285

Pro Val Thr Ser Cys Ser Pro Gly Gly Ala Leu Gln Pro Arg Arg Leu
 290 295 300

Gly Arg Gly Pro Gly Asn Pro Arg Thr Arg Val Cys Gly Gly Gly Ile
 305 310 315 320

Cys Val Val Ala Leu Ala Leu Gln Arg Leu Val Arg Ala Val Arg Arg
 325 330 335

Arg Glu Gly Ala Ala Leu Gly Leu Val Ser Leu Val Val Val Arg Pro
 340 345 350

Val Gly Ala Leu Pro Cys Val Leu Arg Val Pro Arg

267

355

360

<210> 198
 <211> 192
 <212> PRT
 <213> Homo sapien

<400> 198

Ala Gln Pro Ala Cys Arg Ala Glu Arg Gly Arg Gly Val Cys Gly Ser
 1 5 10 15

Gln Ala Gly Pro Pro Thr Gly Gly Ser Ser Ala Gln Pro Pro Pro Leu
 20 25 30

Arg Thr Gln Pro Gln His Ser Gly Cys Thr Asp His Ala Cys Ala Val
 35 40 45

Pro Ser Phe Ser Gln Gly Pro Leu Lys Asp Ala Pro Asn Leu Ile Cys
 50 55 60

Thr Pro His Ala Ala Trp Tyr Ser Glu Gln Ala Ser Ile Glu Met Arg
 65 70 75 80

Glu Glu Ala Ala Arg Glu Ile Arg Arg Ala Ile Thr Gly Arg Ile Pro
 85 90 95

Asp Ser Leu Lys Asn Cys Val Asn Lys Asp His Leu Thr Ala Ala Thr
 100 105 110

His Trp Ala Ser Met Asp Pro Ala Val Val His Pro Glu Leu Asn Gly
 115 120 125

Ala Ala Tyr Arg Tyr Pro Pro Gly Val Val Gly Val Ala Pro Thr Gly
 130 135 140

Ile Pro Ala Ala Val Glu Gly Ile Val Pro Ser Ala Met Ser Leu Ser
 145 150 155 160

His Gly Leu Pro Pro Val Ala His Pro Pro His Ala Pro Ser Pro Gly
 165 170 175

Gln Thr Val Lys Pro Glu Ala Asp Arg Asp His Ala Ser Asp Gln Leu
 180 185 190

<210> 199
 <211> 178
 <212> PRT

268

<213> Homo sapien

<400> 199

Met Arg Glu Glu Ala Pro Phe Ser Phe Ser Gln Gly Pro Leu Lys Asp
 1 5 10 15

Ala Pro Asn Leu Ile Cys Thr Pro His Ala Ala Trp Tyr Met Asp Pro
 20 25 30

Ala Val Val His Pro Glu Leu Asn Gly Ala Ala Tyr Ser Arg Gly Thr
 35 40 45

Leu Arg Ala Trp Trp Ala Trp Pro Pro Leu Ala Ser Gln Leu Leu Trp
 50 55 60

Lys Val Ser Ser Pro Ala Pro Cys Pro Cys Pro Thr Ala Cys Pro Leu
 65 70 75 80

Trp Pro Thr Arg Pro Thr Pro Leu Leu Leu Ala Lys Pro Ser Ser Pro
 85 90 95

Arg Arg Ile Glu Thr Thr Pro Val Thr Ser Cys Ser Pro Gly Gly Ala
 100 105 110

Leu Gln Pro Arg Arg Leu Gly Arg Gly Pro Gly Asn Pro Arg Thr Arg
 115 120 125

Val Cys Gly Gly Gly Ile Cys Val Val Ala Leu Ala Leu Gln Arg Leu
 130 135 140

Val Arg Ala Val Arg Arg Arg Glu Gly Ala Ala Leu Gly Leu Val Ser
 145 150 155 160

Leu Val Val Val Arg Pro Val Gly Ala Leu Pro Cys Val Leu Arg Val
 165 170 175

Pro Arg

<210> 200

<211> 162

<212> PRT

<213> Homo sapien

<400> 200

Arg Met His Pro Thr Ser Ser Ala Pro Pro Met Leu His Gly Thr Trp
 1 5 10 15

269

Thr Pro Pro Ser Cys Thr Leu Ser Ser Met Gly Leu Pro Ile Gly Thr
 20 25 30

Leu Arg Ala Trp Trp Ala Trp Pro Pro Leu Ala Ser Gln Leu Leu Trp
 35 40 45

Lys Val Ser Ser Pro Ala Pro Cys Pro Cys Pro Thr Ala Cys Pro Leu
 50 55 60

Trp Pro Thr Arg Pro Thr Pro Leu Leu Leu Ala Lys Pro Ser Ser Pro
 65 70 75 80

Arg Arg Ile Glu Thr Thr Pro Val Thr Ser Cys Ser Pro Gly Gly Ala
 85 90 95

Leu Gln Pro Arg Arg Leu Gly Arg Gly Pro Gly Asn Pro Arg Thr Arg
 100 105 110

Val Cys Gly Gly Gly Ile Cys Val Val Ala Leu Ala Leu Gln Arg Leu
 115 120 125

Val Arg Ala Val Arg Arg Arg Glu Gly Ala Ala Leu Gly Leu Val Ser
 130 135 140

Leu Val Val Val Arg Pro Val Gly Ala Leu Pro Cys Val Leu Arg Val
 145 150 155 160

Pro Arg

<210> 201

<211> 272

<212> PRT

<213> Homo sapien

<400> 201

Ala Ser Cys Gly Val Gly Arg Leu Val Gly Trp Gly Ile Ser Gly Gly
 1 5 10 15

Gly Ala Ser Leu Gly Pro Gly His Leu Gly Gly Gly Ala Ser Trp Gly
 20 25 30

Arg Gly Ile Ser Glu Gly Ala Ser Gly Gly Trp Ser Ile Leu Gly Gly
 35 40 45

270

Gly Ser Arg Trp Gln Arg Gly Phe Pro Gln Leu Ala Gly Gly Val Ile
 50 55 60

Leu Gly Val Ala Leu Trp Leu Arg His Asp Pro Gln Thr Thr Asn Leu
 65 70 75 80

Leu Tyr Leu Glu Leu Gly Asp Lys Pro Ala Pro Asn Thr Phe Tyr Val
 85 90 95

Gly Ile Tyr Ile Leu Ile Ala Val Gly Ala Val Met Met Phe Val Gly
 100 105 110

Phe Leu Gly Cys Tyr Gly Ala Ile Gln Glu Ser Gln Cys Leu Leu Gly
 115 120 125

Thr Phe Phe Thr Cys Leu Val Ile Leu Phe Ala Cys Glu Val Ala Ala
 130 135 140

Gly Ile Trp Gly Phe Val Asn Lys Asp Gln Ile Ala Lys Asp Val Lys
 145 150 155 160

Gln Phe Tyr Asp Gln Ala Leu Gln Gln Ala Val Val Asp Asp Asp Ala
 165 170 175

Asn Asn Ala Lys Ala Val Val Lys Thr Phe His Glu Thr Leu Asp Cys
 180 185 190

Cys Gly Ser Ser Thr Leu Thr Ala Leu Thr Thr Ser Val Leu Lys Asn
 195 200 205

Asn Leu Cys Pro Ser Gly Ser Asn Ile Ile Ser Asn Leu Phe Lys Glu
 210 215 220

Asp Cys His Gln Lys Ile Asp Asp Leu Phe Ser Gly Lys Leu Tyr Leu
 225 230 235 240

Ile Gly Ile Ala Ala Ile Val Val Ala Val Ile Met Ile Phe Glu Met
 245 250 255

Ile Leu Ser Met Val Leu Cys Cys Gly Ile Arg Asn Ser Ser Val Tyr
 260 265 270

<210> 202
 <211> 303
 <212> PRT
 <213> Homo sapien

271

<400> 202

Met Ser Gly Ala Val Thr Ser His Leu Pro Gln Ala Gly Leu Phe Cys
 1 5 10 15

Thr Ala Cys Leu Gly Arg Trp Trp Glu Ser Leu Trp Pro Ser Ala Leu
 20 25 30

Pro Trp Gln Trp Gly Gln Leu Gly His Leu Gly Gly Ala Arg Leu Pro
 35 40 45

Gln Ala Arg Pro Trp Asp Leu Ser Arg Cys Leu Val Val Ala Cys Phe
 50 55 60

Ser Pro Gly Met Trp Glu Arg His Gln Thr Gln Asp Val Pro Leu Pro
 65 70 75 80

Ala Pro Glu Ala Pro Ser Pro Asp Glu Leu Ala Gly Gly Val Ile Leu
 85 90 95

Gly Val Ala Leu Trp Leu Arg His Asp Pro Gln Thr Thr Asn Leu Leu
 100 105 110

Tyr Leu Glu Leu Gly Asp Lys Pro Ala Pro Asn Thr Phe Tyr Val Gly
 115 120 125

Ile Tyr Ile Leu Ile Ala Val Gly Ala Val Met Met Phe Val Gly Phe
 130 135 140

Leu Gly Cys Tyr Gly Ala Ile Gln Glu Ser Gln Cys Leu Leu Gly Thr
 145 150 155 160

Phe Phe Thr Cys Leu Val Ile Leu Phe Ala Cys Glu Val Ala Ala Gly
 165 170 175

Ile Trp Gly Phe Val Asn Lys Asp Gln Ile Ala Lys Asp Val Lys Gln
 180 185 190

Phe Tyr Asp Gln Ala Leu Gln Gln Ala Val Val Asp Asp Asp Ala Asn
 195 200 205

Asn Ala Lys Ala Val Val Lys Thr Phe His Glu Thr Leu Asp Cys Cys
 210 215 220

Gly Ser Ser Thr Leu Thr Ala Leu Thr Thr Ser Val Leu Lys Asn Asn
 225 230 235 240

272

Leu Cys Pro Ser Gly Ser Asn Ile Ile Ser Asn Leu Phe Lys Glu Asp
 245 250 255

Cys His Gln Lys Ile Asp Asp Leu Phe Ser Gly Lys Leu Tyr Leu Ile
 260 265 270

Gly Ile Ala Ala Ile Val Val Ala Val Ile Met Ile Phe Glu Met Ile
 275 280 285

Leu Ser Met Val Leu Cys Cys Gly Ile Arg Asn Ser Ser Val Tyr
 290 295 300

<210> 203
 <211> 420
 <212> PRT
 <213> Homo sapien

<400> 203

Met Leu Pro Ser Gln Gly Ala Trp Gly Ser Ser Gly Gly Leu Ala Tyr
 1 5 10 15

Thr Pro Trp Ser Ser Cys Pro Arg Trp Gly Ala Gly Leu Gln Pro Ser
 20 25 30

Ala Gln Gly Leu Gly Ile Gln Leu Asp Pro Pro His Thr Ala Ala Arg
 35 40 45

Phe Lys Cys Arg Ser Arg Asn Gly Ser Ala Ala Val Gln Pro Arg Leu
 50 55 60

Gly Gly Arg Ser Gln Gln Gly Pro Pro Thr Leu Phe Ser His His Thr
 65 70 75 80

Gly Glu Ala Ala Leu Val Pro Val Pro Val Pro Gly Leu Pro Ser Gln
 85 90 95

Pro Arg Pro Thr Val Gly Pro Thr Leu Cys Leu Leu Met Pro Leu Pro
 100 105 110

Pro His Ala Lys Ser Gln Arg Leu Trp Glu Arg Val Lys Ala Val Gly
 115 120 125

Gly Gly Trp Gln Val Gln Ala Val Gly Gly Gly Cys Gly Arg Trp Arg
 130 135 140

Ala Pro Pro Gln Val Ser Ser Cys Glu Ala Pro Val Ala Ser Thr Gln

273

145		150		155		160									
Ser	Ala	His	Glu	Val	Pro	Ser	Pro	His	Val	Ala	Ser	Leu	Val	Ser	Val
				165					170					175	
Cys	Val	Met	Glu	Glu	Val	Thr	Glu	Ala	Gln	Lys	Thr	His	Gln	Ala	Arg
			180					185					190		
Leu	Gly	Cys	Glu	Val	Pro	Cys	Cys	Ser	Ser	Leu	Ala	Val	Ser	Asn	Pro
		195					200						205		
Thr	Ser	Ser	Gln	Leu	Gly	Gly	Pro	Trp	Trp	Val	Arg	His	Pro	Gly	Pro
	210					215					220				
Ser	Gly	Val	Leu	Gly	Cys	Gly	Glu	Cys	Val	Gly	Thr	His	Leu	Val	Ser
225					230					235					240
Leu	Ser	Pro	Gln	Gly	Ile	Tyr	Ile	Leu	Ile	Ala	Val	Gly	Ala	Val	Met
			245					250						255	
Met	Phe	Val	Gly	Phe	Leu	Gly	Cys	Tyr	Gly	Ala	Ile	Gln	Glu	Ser	Gln
		260						265					270		
Cys	Leu	Leu	Gly	Thr	Phe	Phe	Thr	Cys	Leu	Val	Ile	Leu	Phe	Ala	Cys
		275					280					285			
Glu	Val	Ala	Ala	Gly	Ile	Trp	Gly	Phe	Val	Asn	Lys	Asp	Gln	Ile	Ala
	290					295					300				
Lys	Asp	Val	Lys	Gln	Phe	Tyr	Asp	Gln	Ala	Leu	Gln	Gln	Ala	Val	Val
305				310					315					320	
Asp	Asp	Asp	Ala	Asn	Asn	Ala	Lys	Ala	Val	Val	Lys	Thr	Phe	His	Glu
			325					330						335	
Thr	Leu	Asp	Cys	Cys	Gly	Ser	Ser	Thr	Leu	Thr	Ala	Leu	Thr	Thr	Ser
		340					345						350		
Val	Leu	Lys	Asn	Asn	Leu	Cys	Pro	Ser	Gly	Ser	Asn	Ile	Ile	Ser	Asn
		355					360					365			
Leu	Phe	Lys	Glu	Asp	Cys	His	Gln	Lys	Ile	Asp	Asp	Leu	Phe	Ser	Gly
	370					375					380				
Lys	Leu	Tyr	Leu	Ile	Gly	Ile	Ala	Ala	Ile	Val	Val	Ala	Val	Ile	Met
385					390				395						400

274

Ile Phe Glu Met Ile Leu Ser Met Val Leu Cys Cys Gly Ile Arg Asn
 405 410 415

Ser Ser Val Tyr
 420

<210> 204
 <211> 247
 <212> PRT
 <213> Homo sapien

<400> 204

Ser Pro Ser Cys Val Met Glu Glu Val Thr Glu Ala Gln Lys Thr His
 1 5 10 15

Gln Ala Arg Leu Gly Cys Glu Val Pro Cys Cys Ser Ser Leu Ala Val
 20 25 30

Ser Asn Pro Thr Ser Ser Gln Leu Gly Gly Pro Trp Trp Val Arg His
 35 40 45

Pro Gly Pro Ser Gly Val Leu Gly Cys Gly Glu Cys Val Gly Thr His
 50 55 60

Leu Val Ser Leu Ser Pro Gln Gly Ile Tyr Ile Leu Ile Ala Val Gly
 65 70 75 80

Ala Val Met Met Phe Val Gly Phe Leu Gly Cys Tyr Gly Ala Ile Gln
 85 90 95

Glu Ser Gln Cys Leu Leu Gly Thr Phe Phe Thr Cys Leu Val Ile Leu
 100 105 110

Phe Ala Cys Glu Val Ala Ala Gly Ile Trp Gly Phe Val Asn Lys Asp
 115 120 125

Gln Ile Ala Lys Asp Val Lys Gln Phe Tyr Asp Gln Ala Leu Gln Gln
 130 135 140

Ala Val Val Asp Asp Asp Ala Asn Asn Ala Lys Ala Val Val Lys Thr
 145 150 155 160

Phe His Glu Thr Leu Asp Cys Cys Gly Ser Ser Thr Leu Thr Ala Leu
 165 170 175

275

Thr Thr Ser Val Leu Lys Asn Asn Leu Cys Pro Ser Gly Ser Asn Ile
 180 185 190

Ile Ser Asn Leu Phe Lys Glu Asp Cys His Gln Lys Ile Asp Asp Leu
 195 200 205

Phe Ser Gly Lys Leu Tyr Leu Ile Gly Ile Ala Ala Ile Val Val Ala
 210 215 220

Val Ile Met Ile Phe Glu Met Ile Leu Ser Met Val Leu Cys Cys Gly
 225 230 235 240

Ile Arg Asn Ser Ser Val Tyr
 245

<210> 205
 <211> 236
 <212> PRT
 <213> Homo sapien

<400> 205

Met Gly Val Glu Gly Cys Thr Lys Cys Ile Lys Tyr Leu Leu Phe Val
 1 5 10 15

Phe Asn Phe Val Phe Trp Leu Ala Gly Gly Val Ile Leu Gly Val Ala
 20 25 30

Leu Trp Leu Arg His Asp Pro Gln Thr Thr Asn Leu Leu Tyr Leu Glu
 35 40 45

Leu Gly Asp Lys Pro Ala Pro Asn Thr Phe Tyr Val Gly Ile Tyr Ile
 50 55 60

Leu Ile Ala Val Gly Ala Val Met Met Phe Val Gly Phe Leu Gly Cys
 65 70 75 80

Tyr Gly Ala Ile Gln Glu Ser Gln Cys Leu Leu Gly Thr Phe Phe Thr
 85 90 95

Cys Leu Val Ile Leu Phe Ala Cys Glu Val Ala Ala Gly Ile Trp Gly
 100 105 110

Phe Val Asn Lys Asp Gln Ile Ala Lys Asp Val Lys Gln Phe Tyr Asp
 115 120 125

Gln Ala Leu Gln Gln Ala Val Val Asp Asp Asp Ala Asn Asn Ala Lys
 130 135 140

276

Ala Val Val Lys Thr Phe His Glu Thr Leu Asp Cys Cys Gly Ser Ser
145 150 155 160

Thr Leu Thr Ala Leu Thr Thr Ser Val Leu Lys Asn Asn Leu Cys Pro
165 170 175

Ser Gly Ser Asn Ile Ile Ser Asn Leu Phe Lys Glu Asp Cys His Gln
180 185 190

Lys Ile Asp Asp Leu Phe Ser Gly Lys Leu Tyr Leu Ile Gly Ile Ala
195 200 205

Ala Ile Val Val Ala Val Ile Met Ile Phe Glu Met Ile Leu Ser Met
210 215 220

Val Leu Cys Cys Gly Ile Arg Asn Ser Ser Val Tyr
225 230 235

<210> 206
<211> 256
<212> PRT
<213> Homo sapien

<400> 206

Met Gly Val Glu Gly Cys Thr Lys Cys Ile Lys Tyr Leu Leu Phe Val
1 5 10 15

Phe Asn Phe Val Phe Trp Leu Ala Gly Gly Val Ile Leu Gly Val Ala
20 25 30

Leu Trp Leu Arg His Asp Pro Gln Thr Thr Asn Leu Leu Tyr Leu Glu
35 40 45

Leu Gly Asp Lys Pro Ala Pro Asn Thr Phe Tyr Val Gly Ile Tyr Ile
50 55 60

Leu Ile Ala Val Gly Ala Val Met Met Phe Val Gly Phe Leu Gly Cys
65 70 75 80

Tyr Gly Ala Ile Gln Glu Ser Gln Cys Leu Leu Gly Thr Phe Phe Thr
85 90 95

Cys Leu Val Ile Leu Phe Ala Cys Glu Val Ala Ala Gly Ile Trp Gly
100 105 110

277

Phe Val Asn Lys Asp Gln Ile Ala Lys Asp Val Lys Gln Phe Tyr Asp
 115 120 125

Gln Ala Leu Gln Gln Ala Val Val Asp Asp Asp Ala Asn Asn Ala Lys
 130 135 140

Ala Val Val Lys Thr Phe His Glu Thr Leu Asp Cys Cys Gly Ser Ser
 145 150 155 160

Thr Leu Thr Ala Leu Thr Thr Ser Val Leu Lys Asn Asn Leu Cys Pro
 165 170 175

Ser Gly Ser Asn Ile Ile Ser Asn Leu Phe Lys Glu Asp Cys His Gln
 180 185 190

Lys Ile Asp Asp Leu Phe Ser Gly Lys Leu Tyr Leu Ile Gly Ile Ala
 195 200 205

Ala Ile Val Val Ala Val Ile Met Ile Phe Glu Met Ile Leu Ser Met
 210 215 220

Val Leu Asn Asp Asn Leu Cys Ile Ile Gly Lys Val Arg Ile Ser Gly
 225 230 235 240

Arg Gln Gly Phe Tyr Pro Asn Gln Gln His Lys Arg Gln Tyr Asn Cys
 245 250 255

<210> 207
 <211> 210
 <212> PRT
 <213> Homo sapien

<400> 207

Met Gly Val Glu Gly Cys Thr Lys Cys Ile Lys Tyr Leu Leu Phe Val
 1 5 10 15

Phe Asn Phe Val Phe Trp Leu Ala Gly Gly Val Ile Leu Gly Val Ala
 20 25 30

Leu Trp Leu Arg His Asp Pro Gln Thr Thr Asn Leu Leu Tyr Leu Glu
 35 40 45

Leu Gly Asp Lys Pro Ala Pro Asn Thr Phe Tyr Val Gly Ile Tyr Ile
 50 55 60

Leu Ile Ala Val Gly Ala Val Met Met Phe Val Gly Phe Leu Gly Cys
 65 70 75 80

278

Tyr Gly Ala Ile Gln Glu Ser Gln Cys Leu Leu Gly Thr Phe Phe Thr
 85 90 95

Cys Leu Val Ile Leu Phe Ala Cys Glu Val Ala Ala Gly Ile Trp Gly
 100 105 110

Phe Val Asn Lys Asp Gln Ile Ala Lys Asp Val Lys Gln Phe Tyr Asp
 115 120 125

Gln Ala Leu Gln Gln Ala Val Val Asp Asp Asp Ala Asn Asn Ala Lys
 130 135 140

Ala Val Val Lys Thr Phe His Glu Thr Leu Asp Cys Cys Gly Ser Ser
 145 150 155 160

Thr Leu Thr Ala Leu Thr Thr Ser Val Leu Lys Asn Asn Leu Cys Pro
 165 170 175

Ser Gly Ser Asn Ile Ile Ser Asn Leu Phe Lys Glu Asp Cys His Gln
 180 185 190

Lys Ile Asp Asp Leu Phe Ser Gly Lys Leu Tyr Leu Ala Ala Thr Thr
 195 200 205

Leu Arg
 210

<210> 208
 <211> 58
 <212> PRT
 <213> Homo sapien

<400> 208

Asn His Ile Glu Pro Leu Lys Ile Gln Trp Leu Asp Val Leu Gln Arg
 1 5 10 15

Glu Pro Arg Pro Phe Pro Lys Leu Arg Ile Leu Arg Lys Val Glu Lys
 20 25 30

Ile Asp Asp Phe Lys Ala Glu Asp Phe Gln Ile Glu Gly Tyr Asn Pro
 35 40 45

His Pro Thr Ile Lys Met Glu Met Ala Val
 50 55

279

<210> 209
 <211> 91
 <212> PRT
 <213> Homo sapien

<400> 209

Lys Phe Ser Gly Ser Met Cys Phe Ser Glu Asn Pro Asp Leu Ser Gln
 1 5 10 15

Ser Ser Gly Phe Phe Glu Lys Leu Arg Lys Leu Met Thr Ser Lys Leu
 20 25 30

Lys Thr Phe Arg Leu Lys Gly Thr Ile Arg Ile Gln Leu Leu Lys Trp
 35 40 45

Lys Trp Leu Phe Arg Val Leu Ser Lys Glu Leu Glu Gly Tyr Cys Gln
 50 55 60

Ser Leu Gly Val Gly Leu Asp Ala Glu Val Lys Val Leu Phe Ala Leu
 65 70 75 80

Lys Glu Lys Gly Thr Arg Ser Lys Ile Cys Pro
 85 90

<210> 210
 <211> 86
 <212> PRT
 <213> Homo sapien

<400> 210

Met Asp Asp Ser Glu Val Glu Ser Thr Ala Ser Ile Leu Ala Ser Val
 1 5 10 15

Lys Glu Gln Glu Ala Gln Phe Glu Lys Leu Thr Arg Ala Leu Glu Glu
 20 25 30

Glu Arg Arg His Val Ser Ala Gln Leu Glu Arg Val Arg Val Ser Pro
 35 40 45

Gln Asp Ala Asn Pro Leu Met Ala Asn Gly Thr Ser Pro Phe Arg Lys
 50 55 60

Lys Cys Lys Lys Lys Ser Ile Phe Ser Ser Arg Val Glu Leu Phe Lys
 65 70 75 80

Glu Ser Lys Ile Ile Ser
 85

280

<210> 211
 <211> 107
 <212> PRT
 <213> Homo sapien

<400> 211

Met Ile Ile Tyr Tyr Met Val His Asn His Val Asp Ala Gln Cys Met
 1 5 10 15

Ile Leu Gln Asn Arg Leu Ser Val Ser Arg Arg Val Leu Arg Gly Met
 20 25 30

Val Met Tyr Thr Ser Lys Asp Arg Tyr Phe Tyr Phe Gly Lys Leu Asp
 35 40 45

Gly Gln Ile Ser Ser Ala Tyr Pro Ser Gln Glu Gly Gln Val Leu Val
 50 55 60

Gly Ile Tyr Gly Gln Tyr Gln Leu Leu Gly Ile Lys Ser Ile Gly Phe
 65 70 75 80

Glu Trp Asn Tyr Pro Leu Glu Glu Pro Thr Thr Glu Pro Pro Val Asn
 85 90 95

Leu Thr Tyr Ser Ala Asn Ser Pro Val Gly Arg
 100 105

<210> 212
 <211> 90
 <212> PRT
 <213> Homo sapien

<400> 212

Tyr Cys Arg Ile Gly Leu Arg Val Ala Arg Val Leu Arg Gly Met Val
 1 5 10 15

Met Tyr Thr Ser Lys Asp Arg Tyr Phe Tyr Phe Gly Lys Leu Asp Gly
 20 25 30

Gln Ile Ser Ser Ala Tyr Pro Ser Gln Glu Gly Gln Val Leu Val Gly
 35 40 45

Ile Tyr Gly Gln Tyr Gln Leu Leu Gly Ile Lys Ser Ile Gly Phe Glu
 50 55 60

Trp Asn Tyr Pro Leu Glu Glu Pro Thr Thr Glu Pro Pro Val Asn Leu
 65 70 75 80

281

Thr Tyr Ser Ala Asn Ser Pro Val Gly Arg
85 90

<210> 213
<211> 193
<212> PRT
<213> Homo sapien

<400> 213

Met Asp Glu Arg Pro Pro Gly Gln Val Thr Gly Glu Ser Pro Gly Met
1 5 10 15

His Arg Pro Glu Ala Met Leu Leu Leu Leu Thr Leu Ala Leu Leu Gly
20 25 30

Gly Pro Thr Trp Ala Gly Lys Met Tyr Gly Pro Gly Gly Gly Lys Tyr
35 40 45

Phe Ser Thr Thr Glu Asp Tyr Asp His Glu Ile Thr Gly Leu Arg Val
50 55 60

Ser Val Gly Leu Leu Leu Val Lys Ser Val Gln Val Lys Leu Gly Asp
65 70 75 80

Ser Trp Asp Val Lys Leu Gly Ala Leu Gly Gly Asn Thr Gln Glu Val
85 90 95

Thr Leu Gln Pro Gly Glu Tyr Ile Thr Lys Val Phe Val Ala Phe Gln
100 105 110

Ala Phe Leu Arg Gly Met Val Met Tyr Thr Ser Lys Asp Arg Tyr Phe
115 120 125

Tyr Phe Gly Lys Leu Asp Gly Gln Ile Ser Ser Ala Tyr Pro Ser Gln
130 135 140

Glu Gly Gln Val Leu Val Gly Ile Tyr Gly Gln Tyr Gln Leu Leu Gly
145 150 155 160

Ile Lys Ser Ile Gly Phe Glu Trp Asn Tyr Pro Leu Glu Glu Pro Thr
165 170 175

Thr Glu Pro Pro Val Asn Leu Thr Tyr Ser Ala Asn Ser Pro Val Gly
180 185 190

282

Arg

<210> 214
 <211> 189
 <212> PRT
 <213> Homo sapien

<400> 214

Ala Ala Ala Arg Ala Gly Gly Glu Ser Pro Gly Met His Arg Pro Glu
 1 5 10 15

Ala Met Leu Leu Leu Leu Thr Leu Ala Leu Leu Gly Gly Pro Thr Trp
 20 25 30

Ala Gly Lys Met Tyr Gly Pro Gly Gly Gly Lys Tyr Phe Ser Thr Thr
 35 40 45

Glu Asp Tyr Asp His Glu Ile Thr Gly Leu Arg Val Ser Val Gly Leu
 50 55 60

Leu Leu Val Lys Ser Val Gln Val Lys Leu Gly Asp Ser Trp Asp Val
 65 70 75 80

Lys Leu Gly Ala Leu Gly Gly Asn Thr Gln Glu Val Thr Leu Gln Pro
 85 90 95

Gly Glu Tyr Ile Thr Lys Val Phe Val Ala Phe Gln Ala Phe Leu Arg
 100 105 110

Gly Met Val Met Tyr Thr Ser Lys Asp Arg Tyr Phe Tyr Phe Gly Lys
 115 120 125

Leu Asp Gly Gln Ile Ser Ser Ala Tyr Pro Ser Gln Glu Gly Gln Val
 130 135 140

Leu Val Gly Ile Tyr Gly Gln Tyr Gln Leu Leu Gly Ile Lys Ser Ile
 145 150 155 160

Gly Phe Glu Trp Asn Tyr Pro Leu Glu Glu Pro Thr Thr Glu Pro Pro
 165 170 175

Val Asn Leu Thr Tyr Ser Ala Asn Ser Pro Val Gly Arg
 180 185

<210> 215
 <211> 202

283

<212> PRT

<213> Homo sapien

<400> 215

Met Asp Arg Pro Pro Gly Arg Trp Arg Val Pro Gly Thr Thr Arg Arg
 1 5 10 15

Pro Val Thr Gly Glu Ser Pro Gly Met His Arg Pro Glu Ala Met Leu
 20 25 30

Leu Leu Leu Thr Leu Ala Leu Leu Gly Gly Pro Thr Trp Ala Gly Lys
 35 40 45

Met Tyr Gly Pro Gly Gly Gly Lys Tyr Phe Ser Thr Thr Glu Asp Tyr
 50 55 60

Asp His Glu Ile Thr Gly Leu Arg Val Ser Val Gly Leu Leu Leu Val
 65 70 75 80

Lys Ser Val Gln Val Lys Leu Gly Asp Ser Trp Asp Val Lys Leu Gly
 85 90 95

Ala Leu Gly Gly Asn Thr Gln Glu Val Thr Leu Gln Pro Gly Glu Tyr
 100 105 110

Ile Thr Lys Val Phe Val Ala Phe Gln Ala Phe Leu Arg Gly Met Val
 115 120 125

Met Tyr Thr Ser Lys Asp Arg Tyr Phe Tyr Phe Gly Lys Leu Asp Gly
 130 135 140

Gln Ile Ser Ser Ala Tyr Pro Ser Gln Glu Gly Gln Val Leu Val Gly
 145 150 155 160

Ile Tyr Gly Gln Tyr Gln Leu Leu Gly Ile Lys Ser Ile Gly Phe Glu
 165 170 175

Trp Asn Tyr Pro Leu Glu Glu Pro Thr Thr Glu Pro Pro Val Asn Leu
 180 185 190

Thr Tyr Ser Ala Asn Ser Pro Val Gly Arg
 195 200

<210> 216

<211> 208

<212> PRT

<213> Homo sapien

284

<400> 216

Cys Arg Ala Ala Gln Cys Asp Gly Ser Ala Ala Gly Gln Val Glu Gly
 1 5 10 15

Ala Arg His Asn Gln Thr Pro Ser His Gly Glu Ser Pro Gly Met His
 20 25 30

Arg Pro Glu Ala Met Leu Leu Leu Leu Thr Leu Ala Leu Leu Gly Gly
 35 40 45

Pro Thr Trp Ala Gly Lys Met Tyr Gly Pro Gly Gly Gly Lys Tyr Phe
 50 55 60

Ser Thr Thr Glu Asp Tyr Asp His Glu Ile Thr Gly Leu Arg Val Ser
 65 70 75 80

Val Gly Leu Leu Leu Val Lys Ser Val Gln Val Lys Leu Gly Asp Ser
 85 90 95

Trp Asp Val Lys Leu Gly Ala Leu Gly Gly Asn Thr Gln Glu Val Thr
 100 105 110

Leu Gln Pro Gly Glu Tyr Ile Thr Lys Val Phe Val Ala Phe Gln Ala
 115 120 125

Phe Leu Arg Gly Met Val Met Tyr Thr Ser Lys Asp Arg Tyr Phe Tyr
 130 135 140

Phe Gly Lys Leu Asp Gly Gln Ile Ser Ser Ala Tyr Pro Ser Gln Glu
 145 150 155 160

Gly Gln Val Leu Val Gly Ile Tyr Gly Gln Tyr Gln Leu Leu Gly Ile
 165 170 175

Lys Ser Ile Gly Phe Glu Trp Asn Tyr Pro Leu Glu Glu Pro Thr Thr
 180 185 190

Glu Pro Pro Val Asn Leu Thr Tyr Ser Ala Asn Ser Pro Val Gly Arg
 195 200 205

<210> 217

<211> 189

<212> PRT

<213> Homo sapien

<400> 217

285

Met His Val Glu Arg Arg Ser Val Met Asp Arg Gly Arg Gly Glu Val
1 5 10 15

Ala Met Leu Leu Leu Leu Thr Leu Ala Leu Leu Gly Gly Pro Thr Trp
20 25 30

Ala Gly Lys Met Tyr Gly Pro Gly Gly Gly Lys Tyr Phe Ser Thr Thr
35 40 45

Glu Asp Tyr Asp His Glu Ile Thr Gly Leu Arg Val Ser Val Gly Leu
50 55 60

Leu Leu Val Lys Ser Val Gln Val Lys Leu Gly Asp Ser Trp Asp Val
65 70 75 80

Lys Leu Gly Ala Leu Gly Gly Asn Thr Gln Glu Val Thr Leu Gln Pro
85 90 95

Gly Glu Tyr Ile Thr Lys Val Phe Val Ala Phe Gln Ala Phe Leu Arg
100 105 110

Gly Met Val Met Tyr Thr Ser Lys Asp Arg Tyr Phe Tyr Phe Gly Lys
115 120 125

Leu Asp Gly Gln Ile Ser Ser Ala Tyr Pro Ser Gln Glu Gly Gln Val
130 135 140

Leu Val Gly Ile Tyr Gly Gln Tyr Gln Leu Leu Gly Ile Lys Ser Ile
145 150 155 160

Gly Phe Glu Trp Asn Tyr Pro Leu Glu Glu Pro Thr Thr Glu Pro Pro
165 170 175

Val Asn Leu Thr Tyr Ser Ala Asn Ser Pro Val Gly Arg
180 185

<210> 218
<211> 171
<212> PRT
<213> Homo sapien

<400> 218

Met Leu Glu Arg Arg Ile Val Asn Gly Ser Pro Gly Gln Val Gln Ser
1 5 10 15

Gln Met Tyr Gly Pro Gly Gly Gly Lys Tyr Phe Ser Thr Thr Glu Asp

286

20	25	30
Tyr Asp His Glu Ile Thr Gly Leu Arg Val Ser Val Gly Leu Leu Leu		
35	40	45
Val Lys Ser Val Gln Val Lys Leu Gly Asp Ser Trp Asp Val Lys Leu		
50	55	60
Gly Ala Leu Gly Gly Asn Thr Gln Glu Val Thr Leu Gln Pro Gly Glu		
65	70	75
Tyr Ile Thr Lys Val Phe Val Ala Phe Gln Ala Phe Leu Arg Gly Met		
85	90	95
Val Met Tyr Thr Ser Lys Asp Arg Tyr Phe Tyr Phe Gly Lys Leu Asp		
100	105	110
Gly Gln Ile Ser Ser Ala Tyr Pro Ser Gln Glu Gly Gln Val Leu Val		
115	120	125
Gly Ile Tyr Gly Gln Tyr Gln Leu Leu Gly Ile Lys Ser Ile Gly Phe		
130	135	140
Glu Trp Asn Tyr Pro Leu Glu Glu Pro Thr Thr Glu Pro Pro Val Asn		
145	150	155
Leu Thr Tyr Ser Ala Asn Ser Pro Val Gly Arg		
165	170	

<210> 219
 <211> 171
 <212> PRT
 <213> Homo sapien

<220>
 <221> MISC_FEATURE
 <222> (6)..(6)
 <223> X=any amino acid

<400> 219

His Ala Arg Ala Ala Xaa Cys Asp Gly Ser Pro Gly Gln Val Gln Ser
1 5 10 15

Gln Met Tyr Gly Pro Gly Gly Gly Lys Tyr Phe Ser Thr Thr Glu Asp
20 25 30

Tyr Asp His Glu Ile Thr Gly Leu Arg Val Ser Val Gly Leu Leu Leu

287

35

40

45

Val Lys Ser Val Gln Val Lys Leu Gly Asp Ser Trp Asp Val Lys Leu
 50 55 60

Gly Ala Leu Gly Gly Asn Thr Gln Glu Val Thr Leu Gln Pro Gly Glu
 65 70 75 80

Tyr Ile Thr Lys Val Phe Val Ala Phe Gln Ala Phe Leu Arg Gly Met
 85 90 95

Val Met Tyr Thr Ser Lys Asp Arg Tyr Phe Tyr Phe Gly Lys Leu Asp
 100 105 110

Gly Gln Ile Ser Ser Ala Tyr Pro Ser Gln Glu Gly Gln Val Leu Val
 115 120 125

Gly Ile Tyr Gly Gln Tyr Gln Leu Leu Gly Ile Lys Ser Ile Gly Phe
 130 135 140

Glu Trp Asn Tyr Pro Leu Glu Glu Pro Thr Thr Glu Pro Pro Val Asn
 145 150 155 160

Leu Thr Tyr Ser Ala Asn Ser Pro Val Gly Arg
 165 170

<210> 220
 <211> 156
 <212> PRT
 <213> Homo sapien

<400> 220

Met Val Leu Asp Ser Leu His Pro Gly Lys Glu Asp Gly Gly Ala Glu
 1 5 10 15

Asp Pro Gly Cys Ala Gly Pro Ser Gln Ile Trp Thr Ser Lys Ala Leu
 20 25 30

Pro Leu Ser Ser Val Gln Val Lys Leu Gly Asp Ser Trp Asp Val Lys
 35 40 45

Leu Gly Ala Leu Gly Gly Asn Thr Gln Glu Val Thr Leu Gln Pro Gly
 50 55 60

Glu Tyr Ile Thr Lys Val Phe Val Ala Phe Gln Ala Phe Leu Arg Gly
 65 70 75 80

288

Met. Val Met Tyr Thr Ser Lys Asp Arg Tyr Phe Tyr Phe Gly Lys Leu
85 90 95

Asp Gly Gln Ile Ser Ser Ala Tyr Pro Ser Gln Glu Gly Gln Val Leu
100 105 110

Val Gly Ile Tyr Gly Gln Tyr Gln Leu Leu Gly Ile Lys Ser Ile Gly
115 120 125

Phe Glu Trp Asn Tyr Pro Leu Glu Glu Pro Thr Thr Glu Pro Pro Val
130 135 140

Asn Leu Thr Tyr Ser Ala Asn Ser Pro Val Gly Arg
145 150 155

<210> 221
<211> 156
<212> PRT
<213> Homo sapien

<400> 221

Trp Cys Trp Thr Leu Cys Ile Pro Gly Arg Arg Met Gly Ala Leu Arg
1 5 10 15

Thr Arg Asp Val Leu Gly His Pro Arg Ser Gly Arg Pro Lys Leu Cys
20 25 30

Leu Ser Pro Ser Val Gln Val Lys Leu Gly Asp Ser Trp Asp Val Lys
35 40 45

Leu Gly Ala Leu Gly Gly Asn Thr Gln Glu Val Thr Leu Gln Pro Gly
50 55 60

Glu Tyr Ile Thr Lys Val Phe Val Ala Phe Gln Ala Phe Leu Arg Gly
65 70 75 80

Met Val Met Tyr Thr Ser Lys Asp Arg Tyr Phe Tyr Phe Gly Lys Leu
85 90 95

Asp Gly Gln Ile Ser Ser Ala Tyr Pro Ser Gln Glu Gly Gln Val Leu
100 105 110

Val Gly Ile Tyr Gly Gln Tyr Gln Leu Leu Gly Ile Lys Ser Ile Gly
115 120 125

Phe Glu Trp Asn Tyr Pro Leu Glu Glu Pro Thr Thr Glu Pro Pro Val

289

130 135 140
 Asn Leu Thr Tyr Ser Ala Asn Ser Pro Val Gly Arg
 145 150 155

 <210> 222
 <211> 76
 <212> PRT
 <213> Homo sapien

 <400> 222
 Met Val Met Tyr Thr Ser Lys Asp Arg Tyr Phe Tyr Phe Gly Lys Leu
 1 5 10 15

 Asp Gly Gln Ile Ser Ser Ala Tyr Pro Ser Gln Glu Gly Gln Val Leu
 20 25 30

 Val Gly Ile Tyr Gly Gln Tyr Gln Leu Leu Gly Ile Lys Ser Ile Gly
 35 40 45

 Phe Glu Trp Asn Tyr Pro Leu Glu Glu Pro Thr Thr Glu Pro Pro Val
 50 55 60

 Asn Leu Thr Tyr Ser Ala Asn Ser Pro Val Gly Arg
 65 70 75

 <210> 223
 <211> 139
 <212> PRT
 <213> Homo sapien

 <400> 223
 Leu Cys Arg Gly Gln Lys Glu Ser Ser Thr Thr Pro Ser Glu Val Leu
 1 5 10 15

 Trp Ile Ser Val Pro Val Pro Gln Ser Leu Lys Ser Gln Ala Ser Arg
 20 25 30

 Pro Pro Leu Pro Thr Val Pro His Pro Arg Pro Thr Gln Arg Ala Ser
 35 40 45

 Ala Gly His Ser Val Pro Gly Phe Ser Glu Cys Ser Arg Gly Leu Arg
 50 55 60

 Glu Ala Thr His Ser Ser Ile His Ser Ala Asn Ile Cys Gln Gly Arg
 65 70 75 80

290

Val Leu Thr Arg Leu Ala Trp His Trp Gly Tyr Lys Glu Glu Ala Arg
 85 90 95

Phe Gln Leu Ser Ala Tyr Thr Leu Trp Trp Gly Leu Val Gln Arg Gln
 100 105 110

Ile Val Ala Val His Phe Ala Ile Cys Met Asp Gly Asp Thr Cys Arg
 115 120 125

Ser Leu Cys Val Gly Thr Cys Pro Glu Val Arg
 130 135

<210> 224

<211> 568

<212> PRT

<213> Homo sapien

<400> 224

Met Val Lys Leu Ala Lys Ala Gly Lys Asn Gln Gly Asp Pro Lys Lys
 1 5 10 15

Met Ala Pro Pro Pro Lys Glu Val Glu Glu Asp Ser Glu Asp Glu Glu
 20 25 30

Met Ser Glu Asp Glu Glu Asp Asp Ser Ser Gly Glu Glu Val Val Ile
 35 40 45

Pro Gln Lys Lys Gly Lys Lys Ala Ala Ala Thr Ser Ala Lys Lys Val
 50 55 60

Val Val Ser Pro Thr Lys Lys Val Ala Val Ala Thr Pro Ala Lys Lys
 65 70 75 80

Ala Ala Val Thr Pro Gly Lys Lys Ala Ala Ala Thr Pro Ala Lys Lys
 85 90 95

Thr Val Thr Pro Ala Lys Ala Val Thr Thr Pro Gly Lys Lys Gly Ala
 100 105 110

Thr Pro Gly Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Ala
 115 120 125

Ile Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp
 130 135 140

Ser Asp Glu Glu Glu Asp Asp Asp Ser Glu Glu Asp Glu Glu Asp Asp
 145 150 155 160

291

Glu Asp Glu Asp Glu Asp Glu Asp Glu Ile Glu Pro Ala Ala Met Lys
 165 170 175

Ala Ala Ala Ala Ala Pro Ala Ser Glu Asp Glu Asp Asp Glu Asp Asp
 180 185 190

Glu Asp Asp Glu Asp Asp Asp Asp Glu Glu Asp Asp Ser Glu Glu
 195 200 205

Glu Ala Met Glu Thr Thr Pro Ala Lys Gly Lys Lys Ala Ala Lys Val
 210 215 220

Val Pro Val Lys Ala Lys Asn Val Ala Glu Asp Glu Asp Glu Glu Glu
 225 230 235 240

Asp Asp Glu Asp Glu Asp Asp Asp Asp Asp Glu Asp Asp Glu Asp Asp
 245 250 255

Asp Asp Glu Asp Asp Glu Glu Glu Glu Glu Glu Glu Glu Glu Pro
 260 265 270

Val Lys Glu Ala Pro Gly Lys Arg Lys Lys Glu Met Ala Lys Gln Lys
 275 280 285

Ala Ala Pro Glu Ala Lys Lys Gln Lys Val Glu Gly Thr Glu Pro Thr
 290 295 300

Thr Ala Phe Asn Leu Phe Val Gly Asn Leu Asn Phe Asn Lys Ser Ala
 305 310 315 320

Pro Glu Leu Lys Thr Gly Ile Ser Asp Val Phe Ala Lys Asn Asp Leu
 325 330 335

Ala Val Val Asp Val Arg Ile Gly Met Thr Arg Lys Phe Gly Tyr Val
 340 345 350

Asp Phe Glu Ser Ala Glu Asp Leu Glu Lys Ala Leu Glu Leu Thr Gly
 355 360 365

Leu Lys Val Phe Gly Asn Glu Ile Lys Leu Glu Lys Pro Lys Gly Lys
 370 375 380

Asp Ser Lys Lys Glu Arg Asp Ala Arg Thr Leu Leu Ala Lys Asn Leu
 385 390 395 400

292

Pro Tyr Lys Val Thr Gln Asp Glu Leu Lys Glu Val Phe Glu Asp Ala
 405 410 415

Ala Glu Ile Arg Leu Val Ser Lys Asp Gly Lys Ser Lys Gly Ile Ala
 420 425 430

Tyr Ile Glu Phe Lys Thr Glu Ala Asp Ala Glu Lys Thr Phe Glu Glu
 435 440 445

Lys Gln Gly Thr Glu Ile Asp Gly Arg Ser Ile Ser Leu Tyr Tyr Thr
 450 455 460

Gly Glu Lys Gly Gln Asn Gln Asp Tyr Arg Gly Gly Lys Asn Ser Thr
 465 470 475 480

Trp Ser Gly Glu Ser Lys Thr Leu Val Leu Ser Asn Leu Ser Tyr Ser
 485 490 495

Ala Thr Glu Glu Thr Leu Gln Glu Val Phe Glu Lys Ala Thr Phe Ile
 500 505 510

Lys Val Pro Gln Asn Gln Asn Gly Lys Ser Lys Gly Tyr Ala Phe Ile
 515 520 525

Glu Phe Ala Ser Phe Glu Asp Ala Lys Glu Ala Leu Asn Ser Cys Asn
 530 535 540

Lys Arg Glu Ile Glu Gly Arg Ala Ile Arg Leu Glu Ala Arg Arg Leu
 545 550 555 560

Pro Arg Arg Gln Arg Arg Arg Arg
 565

<210> 225
 <211> 520
 <212> PRT
 <213> Homo sapien

<400> 225

Met Val Lys Leu Ala Lys Ala Gly Lys Asn Gln Gly Asp Pro Lys Lys
 1 5 10 15

Met Ala Pro Pro Pro Lys Glu Val Glu Glu Asp Ser Glu Asp Glu Glu
 20 25 30

Met Ser Glu Asp Glu Glu Asp Asp Ser Ser Gly Glu Glu Val Val Ile

293

35

40

45

Pro Gln Lys Lys Gly Lys Lys Ala Ala Ala Thr Ser Ala Lys Lys Val
 50 55 60

Val Val Ser Pro Thr Lys Lys Val Ala Val Ala Thr Pro Ala Lys Lys
 65 70 75 80

Ala Ala Val Thr Pro Gly Lys Lys Ala Ala Ala Thr Pro Ala Lys Lys
 85 90 95

Thr Val Thr Pro Ala Lys Ala Val Thr Thr Pro Gly Lys Lys Gly Ala
 100 105 110

Thr Pro Gly Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Ala
 115 120 125

Ile Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp
 130 135 140

Ser Asp Glu Glu Glu Asp Asp Asp Ser Glu Glu Asp Glu Glu Asp Asp
 145 150 155 160

Glu Asp Glu Asp Glu Asp Glu Asp Glu Ile Glu Pro Ala Ala Met Lys
 165 170 175

Ala Ala Ala Ala Ala Pro Ala Ser Glu Asp Glu Asp Asp Glu Asp Asp
 180 185 190

Glu Asp Asp Glu Asp Asp Asp Asp Asp Glu Glu Asp Asp Ser Glu Glu
 195 200 205

Glu Ala Met Glu Thr Thr Pro Ala Lys Gly Lys Lys Ala Ala Lys Val
 210 215 220

Val Pro Val Lys Ala Lys Asn Val Ala Glu Asp Glu Asp Glu Glu Glu
 225 230 235 240

Asp Asp Glu Asp Glu Asp Asp Asp Asp Asp Glu Asp Asp Glu Asp Asp
 245 250 255

Asp Asp Glu Asp Asp Glu Glu Glu Glu Glu Glu Glu Glu Glu Pro
 260 265 270

Val Lys Glu Ala Pro Gly Lys Arg Lys Lys Glu Met Ala Lys Gln Lys
 275 280 285

294

Ala Ala Pro Glu Ala Lys Lys Gln Lys Val Glu Gly Thr Glu Pro Thr
 290 295 300

Thr Ala Phe Asn Leu Phe Val Gly Asn Leu Asn Phe Asn Lys Ser Ala
 305 310 315 320

Pro Glu Leu Lys Thr Gly Ile Ser Asp Val Phe Ala Lys Asn Asp Leu
 325 330 335

Ala Val Val Asp Val Arg Ile Gly Met Thr Arg Lys Phe Gly Tyr Val
 340 345 350

Asp Phe Glu Ser Ala Glu Asp Leu Glu Lys Ala Leu Glu Leu Thr Gly
 355 360 365

Leu Lys Val Phe Gly Asn Glu Ile Lys Leu Glu Lys Pro Lys Gly Lys
 370 375 380

Asp Ser Lys Lys Glu Arg Asp Ala Arg Thr Leu Leu Ala Lys Asn Leu
 385 390 395 400

Pro Tyr Lys Val Thr Gln Asp Glu Leu Lys Glu Val Phe Glu Asp Ala
 405 410 415

Ala Glu Ile Arg Leu Val Ser Lys Asp Gly Lys Ser Lys Gly Ile Ala
 420 425 430

Tyr Ile Glu Phe Lys Thr Glu Ala Asp Ala Glu Lys Thr Phe Glu Glu
 435 440 445

Lys Gln Gly Thr Glu Ile Asp Gly Arg Ser Ile Ser Leu Tyr Tyr Thr
 450 455 460

Gly Glu Lys Gly Gln Asn Gln Asp Tyr Arg Gly Gly Lys Asn Ser Thr
 465 470 475 480

Trp Ser Gly Glu Ser Lys Thr Leu Val Leu Ser Asn Leu Ser Tyr Ser
 485 490 495

Ala Thr Glu Glu Thr Leu Gln Glu Val Phe Glu Lys Ala Thr Phe Ile
 500 505 510

Lys Val Pro Arg Pro Arg Pro Arg
 515 520

295

<210> 226
 <211> 526
 <212> PRT
 <213> Homo sapien

<400> 226

Met Leu Arg Leu Pro Thr Val Phe Arg Gln Met Arg Pro Val Ser Arg
 1 5 10 15

Val Leu Ala Pro His Leu Thr Arg Ala Tyr Ala Lys Asp Val Lys Phe
 20 25 30

Gly Ala Asp Ala Arg Ala Leu Met Leu Gln Gly Val Asp Leu Leu Ala
 35 40 45

Asp Ala Val Ala Val Thr Met Gly Pro Lys Gly Arg Thr Val Ile Ile
 50 55 60

Glu Gln Ser Trp Gly Ser Pro Lys Val Thr Lys Asp Gly Val Thr Val
 65 70 75 80

Ala Lys Ser Ile Asp Leu Lys Asp Lys Tyr Lys Asn Ile Gly Ala Lys
 85 90 95

Leu Val Gln Asp Val Ala Asn Asn Thr Asn Glu Glu Ala Gly Asp Gly
 100 105 110

Thr Thr Thr Ala Thr Val Leu Ala Arg Ser Ile Ala Lys Glu Gly Phe
 115 120 125

Glu Lys Ile Ser Lys Gly Ala Asn Pro Val Glu Ile Arg Arg Gly Val
 130 135 140

Met Leu Ala Val Asp Ala Val Ile Ala Glu Leu Lys Lys Gln Ser Lys
 145 150 155 160

Pro Val Thr Thr Pro Glu Glu Ile Ala Gln Val Ala Thr Ile Ser Ala
 165 170 175

Asn Gly Asp Lys Glu Ile Gly Asn Ile Ile Ser Asp Ala Met Lys Lys
 180 185 190

Val Gly Arg Lys Gly Val Ile Thr Val Lys Asp Gly Lys Thr Leu Asn
 195 200 205

Asp Glu Leu Glu Ile Ile Glu Gly Met Lys Phe Asp Arg Gly Tyr Ile

296

210	215	220
Ser Pro Tyr Phe Ile Asn Thr Ser Lys Gly Gln Lys Cys Glu Phe Gln		
225	230	235 240
Asp Ala Tyr Val Leu Leu Ser Glu Lys Lys Ile Ser Ser Ile Gln Ser		
	245	250 255
Ile Val Pro Ala Leu Glu Ile Ala Asn Ala His Arg Lys Pro Leu Val		
	260	265 270
Ile Ile Ala Glu Asp Val Asp Gly Glu Ala Leu Ser Thr Leu Val Leu		
	275	280 285
Asn Arg Leu Lys Val Gly Leu Gln Val Val Ala Val Lys Ala Pro Gly		
	290	295 300
Phe Gly Asp Asn Arg Lys Asn Gln Leu Lys Asp Met Ala Ile Ala Thr		
305	310	315 320
Gly Gly Ala Val Phe Gly Glu Glu Gly Leu Thr Leu Asn Leu Glu Asp		
	325	330 335
Val Gln Pro His Asp Leu Gly Lys Val Gly Glu Val Ile Val Thr Lys		
	340	345 350
Asp Asp Ala Met Leu Leu Lys Gly Lys Gly Asp Lys Ala Gln Ile Glu		
	355	360 365
Lys Arg Ile Gln Glu Ile Ile Glu Gln Leu Asp Val Thr Thr Ser Glu		
	370	375 380
Tyr Glu Lys Glu Lys Leu Asn Glu Arg Leu Ala Lys Leu Ser Asp Gly		
385	390	395 400
Val Ala Val Leu Lys Val Gly Gly Thr Ser Asp Val Glu Val Asn Glu		
	405	410 415
Lys Lys Asp Arg Val Thr Asp Ala Leu Asn Ala Thr Arg Ala Ala Val		
	420	425 430
Glu Glu Gly Ile Val Leu Gly Gly Gly Cys Ala Leu Leu Arg Cys Ile		
	435	440 445
Pro Ala Leu Asp Ser Leu Thr Pro Ala Asn Glu Asp Gln Lys Ile Gly		
	450	455 460

297

Ile Glu Ile Ile Lys Arg Thr Leu Lys Ile Pro Ala Met Thr Ile Ala
 465 470 475 480

Lys Asn Ala Gly Val Glu Gly Ser Leu Ile Val Glu Lys Ile Met Gln
 485 490 495

Ser Ser Ser Glu Val Gly Tyr Asp Ala Met Ala Gly Asp Phe Val Asn
 500 505 510

Met Val Glu Lys Gly Ile Ile Asp Pro Thr Lys Val Asn Gly
 515 520 525

<210> 227
 <211> 121
 <212> PRT
 <213> Homo sapien

<400> 227

Gln Cys Asp Gly Phe Ala Ala Glu Val Ser Thr Val His Glu Ile Leu
 1 5 10 15

Cys Lys Leu Ser Leu Glu Gly Asp His Ser Thr Pro Pro Ser Ala Tyr
 20 25 30

Gly Ser Val Lys Ala Tyr Thr Asn Phe Asp Ala Glu Arg Asp Ala Leu
 35 40 45

Asn Ile Glu Thr Ala Ile Lys Thr Lys Glu Ala Val Asp Glu Val Thr
 50 55 60

Ile Val Asn Ile Leu Thr Asn Arg Ser Asn Ala Gln Arg Gln Asp Ile
 65 70 75 80

Ala Phe Ala Tyr Gln Arg Arg Thr Lys Lys Glu Leu Ala Ser Ala Leu
 85 90 95

Lys Ser Ala Leu Ser Gly His Leu Glu Thr Val Ile Leu Gly Leu Leu
 100 105 110

Lys Thr Pro Ala Gln Tyr Asp Ala Ser
 115 120

<210> 228
 <211> 71
 <212> PRT
 <213> Homo sapien

298

<400> 228

Asn Ser His Gln Asp Gln Arg Gly Val Asp Glu Val Thr Ile Val Asn
 1 5 10 15

Ile Leu Thr Asn Arg Ser Asn Ala Gln Arg Gln Asp Ile Ala Phe Ala
 20 25 30

Tyr Gln Arg Arg Thr Lys Lys Glu Leu Ala Ser Ala Leu Lys Ser Ala
 35 40 45

Leu Ser Gly His Leu Glu Thr Val Ile Leu Gly Leu Leu Lys Thr Pro
 50 55 60

Ala Gln Tyr Asp Ala Ser Glu
 65 70

<210> 229

<211> 242

<212> PRT

<213> Homo sapien

<400> 229

Met Leu Glu Arg Arg Ser Val Met Asp Val Val Ala Ala Glu Gly Arg
 1 5 10 15

Ser Gln Leu Ser Ala His Gly Pro Ala Ser Phe Lys Met Ser Thr Val
 20 25 30

His Glu Ile Leu Cys Lys Leu Ser Leu Glu Gly Asp His Ser Thr Pro
 35 40 45

Pro Ser Ala Tyr Gly Ser Val Lys Ala Tyr Thr Asn Phe Asp Ala Glu
 50 55 60

Arg Asp Ala Leu Asn Ile Glu Thr Ala Ile Lys Thr Lys Gly Val Asp
 65 70 75 80

Glu Val Thr Ile Val Asn Ile Leu Thr Asn Arg Ser Asn Ala Gln Arg
 85 90 95

Gln Asp Ile Ala Phe Ala Tyr Gln Arg Arg Thr Lys Lys Glu Leu Ala
 100 105 110

Ser Ala Leu Lys Ser Ala Leu Ser Gly His Leu Glu Thr Val Ile Leu
 115 120 125

299

Gly Leu Leu Lys Thr Pro Ala Gln Tyr Asp Ala Ser Glu Leu Cys Ser
 130 135 140

Arg Thr Asn Gln Glu Leu Gln Glu Ile Asn Arg Val Tyr Lys Glu Met
 145 150 155 160

Tyr Lys Thr Asp Leu Glu Lys Asp Ile Ile Ser Asp Thr Ser Gly Asp
 165 170 175

Phe Arg Lys Leu Met Val Ala Leu Ala Lys Gly Arg Arg Ala Glu Asp
 180 185 190

Gly Ser Val Ile Asp Tyr Glu Leu Ile Asp Gln Asp Ala Arg Asp Leu
 195 200 205

Tyr Asp Ala Gly Val Lys Arg Val Lys Arg Lys Gly Thr Asp Val Pro
 210 215 220

Lys Trp Ile Ser Ile Met Thr Glu Arg Ser Val Ala Pro Pro Pro Glu
 225 230 235 240

Ser Ile

<210> 230
 <211> 342
 <212> PRT
 <213> Homo sapien

<400> 230

Trp Ile Val Val Ala Ala Glu Gly Arg Ser Gln Leu Ser Ala His Gly
 1 5 10 15

Pro Ala Ser Phe Lys Met Ser Thr Val His Glu Ile Leu Cys Lys Leu
 20 25 30

Ser Leu Glu Gly Asp His Ser Thr Pro Pro Ser Ala Tyr Gly Ser Val
 35 40 45

Lys Ala Tyr Thr Asn Phe Asp Ala Glu Arg Asp Ala Leu Asn Ile Glu
 50 55 60

Thr Ala Ile Lys Thr Lys Gly Val Asp Glu Val Thr Ile Val Asn Ile
 65 70 75 80

Leu Thr Asn Arg Ser Asn Ala Gln Arg Gln Asp Ile Ala Phe Ala Tyr

300

85	90	95
Gln Arg Arg Thr Lys Lys Glu Leu Ala Ser Ala Leu Lys Ser Ala Leu 100 105 110		
Ser Gly His Leu Glu Thr Val Ile Leu Gly Leu Leu Lys Thr Pro Ala 115 120 125		
Gln Tyr Asp Ala Ser Glu Leu Cys Ser Arg Thr Asn Gln Glu Leu Gln 130 135 140		
Glu Ile Asn Arg Val Tyr Lys Glu Met Tyr Lys Thr Asp Leu Glu Lys 145 150 155 160		
Asp Ile Ile Ser Asp Thr Ser Gly Asp Phe Arg Lys Leu Met Val Ala 165 170 175		
Leu Ala Lys Gly Arg Arg Ala Glu Asp Gly Ser Val Ile Asp Tyr Glu 180 185 190		
Leu Ile Asp Gln Asp Ala Arg Asp Leu Tyr Asp Ala Gly Val Lys Arg 195 200 205		
Lys Gly Thr Asp Val Pro Lys Trp Ile Ser Ile Met Thr Glu Arg Ser 210 215 220		
Val Pro His Leu Gln Lys Val Phe Asp Arg Tyr Lys Ser Tyr Ser Pro 225 230 235 240		
Tyr Asp Met Leu Glu Ser Ile Arg Lys Glu Val Lys Gly Asp Leu Glu 245 250 255		
Asn Ala Phe Leu Asn Leu Val Gln Cys Ile Gln Asn Lys Pro Leu Tyr 260 265 270		
Phe Ala Asp Arg Leu Tyr Asp Ser Met Lys Gly Lys Gly Thr Arg Asp 275 280 285		
Lys Val Leu Ile Arg Ile Met Val Ser Arg Ser Glu Val Asp Met Leu 290 295 300		
Lys Ile Arg Ser Glu Phe Lys Arg Lys Tyr Gly Lys Ser Leu Tyr Tyr 305 310 315 320		
Tyr Ile Gln Gln Asp Thr Lys Gly Asp Tyr Gln Lys Ala Leu Leu Tyr 325 330 335		

301

Leu Cys Gly Gly Asp Asp
340

<210> 231
<211> 72
<212> PRT
<213> Homo sapien

<400> 231

Pro Arg Pro Leu Leu Ala Arg Arg Tyr Leu Cys Arg Val Thr Ser Cys
1 5 10 15

Phe Leu Ser Leu Ser Arg Ala Val Trp Trp Gln Gln Ala Gln Pro Gln
20 25 30

Ala Gln Ala Gln Pro Arg Asn Ala Glu Arg Arg Arg Arg Val Arg Gly
35 40 45

Pro Val Arg Ala Ala Glu Met Arg Pro Leu Ala Ile Ala Ser Ser Val
50 55 60

Pro Arg Thr Thr His Pro Ser Arg
65 70

<210> 232
<211> 103
<212> PRT
<213> Homo sapien

<400> 232

Leu Leu Pro Phe Ser Leu Ala Arg Gly Val Val Ala Ala Gly Ala Ala
1 5 10 15

Gly Ala Pro Ser Leu Glu Met Gln Asn Asp Ala Gly Glu Phe Val Asp
20 25 30

Leu Tyr Val Pro Arg Lys Cys Ser Ala Ser Asn Arg Ile Ile Gly Ala
35 40 45

Lys Asp His Ala Ser Ile Gln Met Asn Val Ala Glu Val Asp Lys Val
50 55 60

Thr Gly Arg Phe Asn Gly Gln Phe Lys Thr Tyr Ala Ile Cys Gly Ala
65 70 75 80

Ile Arg Arg Met Gly Glu Ser Asp Asp Ser Ile Leu Arg Leu Ala Lys

302

85

90

95

Ala Asp Gly Ile Val Ser Lys
100

<210> 233
<211> 112
<212> PRT
<213> Homo sapien

<400> 233

Leu Leu Pro Phe Ser Leu Ala Arg Gly Val Val Ala Ala Gly Ala Ala
1 5 10 15

Gly Ala Pro Ser Leu Glu Met Gln Asn Asp Ala Gly Glu Phe Val Asp
20 25 30

Leu Tyr Val Pro Arg Lys Cys Ser Ala Ser Asn Arg Ile Ile Gly Ala
35 40 45

Lys Asp His Ala Ser Ile Gln Met Asn Val Ala Glu Val Asp Lys Val
50 55 60

Thr Gly Arg Phe Asn Gly Gln Phe Lys Thr Tyr Ala Ile Cys Gly Ala
65 70 75 80

Ile Arg Arg Met Val Ser Val Ser Leu Gly Phe Ala His His Phe Gly
85 90 95

Thr Ser Trp Thr Leu Pro Cys Ala Leu Glu Cys Val Met Val Pro Glu
100 105 110

<210> 234
<211> 87
<212> PRT
<213> Homo sapien

<400> 234

Ala Arg Gly Ile Ala Arg Gly Val Val Ala Ala Gly Ala Ala Gly Ala
1 5 10 15

Gly Pro Ala Ser Lys Cys Arg Thr Thr Pro Ala Ser Ser Trp Thr Cys
20 25 30

Thr Cys Arg Gly Asn Ala Ser Ala Ser Asn Arg Ile Ile Gly Ala Lys
35 40 45

303

Asp His Ala Ser Ile Gln Met Asn Val Ala Glu Val Ser Trp Glu Pro
 50 55 60

Gly Arg Arg Glu Gly Cys Asp Ile Cys Ala Gly Lys Ala Gly Cys Pro
 65 70 75 80

Ile Val Glu Glu Pro Leu Gly
 85

<210> 235
 <211> 86
 <212> PRT
 <213> Homo sapien

<400> 235

Ala Arg Gly Ile Ala Arg Gly Val Val Ala Ala Gly Ala Ala Gly Ala
 1 5 10 15

Pro Ser Leu Glu Met Gln Asn Asp Ala Gly Glu Phe Val Asp Leu Tyr
 20 25 30

Val Pro Arg Lys Cys Ser Ala Ser Asn Arg Ile Ile Gly Ala Lys Asp
 35 40 45

His Ala Ser Ile Gln Met Asn Val Ala Glu Val Ser Trp Glu Pro Gly
 50 55 60

Arg Arg Glu Gly Cys Asp Ile Cys Ala Gly Lys Ala Gly Cys Pro Ile
 65 70 75 80

Val Glu Glu Pro Leu Gly
 85

<210> 236
 <211> 77
 <212> PRT
 <213> Homo sapien

<400> 236

Met Arg Gly Arg Gly Arg Gly Thr Cys Arg Gly Asn Ala Ser Ala Ser
 1 5 10 15

Asn Arg Ile Ile Gly Ala Lys Asp His Ala Ser Ile Gln Met Asn Val
 20 25 30

Ala Glu Val Asp Lys Val Thr Gly Arg Phe Asn Gly Gln Phe Lys Thr
 35 40 45

304

Tyr Ala Ile Cys Gly Ala Ile Arg Arg Met Gly Glu Ser Asp Asp Ser
50 55 60

Ile Leu Arg Leu Ala Lys Ala Asp Gly Ile Val Ser Lys
65 70 75

<210> 237
<211> 86
<212> PRT
<213> Homo sapien

<400> 237

Ile Met Pro Ser Gly Ala Ser Val Met Asp Ala Trp Ser Arg Pro Arg
1 5 10 15

Tyr Val Pro Arg Lys Cys Ser Ala Ser Asn Arg Ile Ile Gly Ala Lys
20 25 30

Asp His Ala Ser Ile Gln Met Asn Val Ala Glu Val Asp Lys Val Thr
35 40 45

Gly Arg Phe Asn Gly Gln Phe Lys Thr Tyr Ala Ile Cys Gly Ala Ile
50 55 60

Arg Arg Met Gly Glu Ser Asp Asp Ser Ile Leu Arg Leu Ala Lys Ala
65 70 75 80

Asp Gly Ile Val Ser Lys
85

<210> 238
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 238
tacgcagagc tcacgtcct t

21

<210> 239
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 239

305

acaaccacga agagccagtc tt

22

<210> 240

<211> 28

<212> DNA

<213> Artificial sequence

<220>

<223> Synthetic

<400> 240

tggtgagct cttacctggt ttccaggc

28